

SANDRINE MICALLEF

FRÉDÉRIC YVES BOIS

**Applications des modèles physiologiques à l'analyse
de données de toxicocinétique**

Journal de la société française de statistique, tome 145, n° 3 (2004),
p. 15-32

http://www.numdam.org/item?id=JSFS_2004__145_3_15_0

© Société française de statistique, 2004, tous droits réservés.

L'accès aux archives de la revue « Journal de la société française de statistique » (<http://publications-sfds.math.cnrs.fr/index.php/J-SFdS>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

APPLICATIONS DES MODÈLES PHYSIOLOGIQUES À L'ANALYSE DE DONNÉES DE TOXICOCINÉTIQUE

Sandrine MICALLEF et Frédéric Yves BOIS *

RÉSUMÉ

Les études de toxicologie visent à protéger les populations humaines en quantifiant l'impact d'expositions aux produits chimiques sur la santé. La toxicocinétique s'attache plus particulièrement à décrire la relation entre une exposition et la quantité résultante de toxique au niveau des tissus du corps. À cette fin, des modèles mathématiques sont utilisés pour comprendre et représenter les phénomènes physiologiques de transport, transformation et élimination des substances chimiques. Pour des raisons éthiques, les données permettant d'estimer les paramètres de ces modèles sont éparses et ne concernent que rarement la population humaine. D'une part, les extrapolations requises (inter-espèces, inter-sujets, etc.) sont facilitées par l'emploi de modèles toxicocinétiques « physiologiques ». Nous verrons d'autre part que le problème posé par la pauvreté des données peut être résolu par leur analyse dans un cadre bayésien. L'estimation bayésienne des paramètres de ces modèles limite le besoin de recourir à l'expérimentation en utilisant l'information disponible *a priori* sur la valeur des paramètres physiologiques (flux sanguin, volume des organes du corps, etc.)

Mots clés : Modèles toxicocinétiques, Statistique bayésienne, Simulations stochastiques, Toxicologie.

ABSTRACT

Toxicological studies aim at protecting human populations by quantifying the damages caused by exposure to chemical agents. The main objective of toxicokinetic studies is to describe the relationships between exposure and the effective quantity of toxic reaching tissues of the body. In that context, mathematical models are used to understand physiological phenomena of transport, transformation and elimination of chemicals in the body. For ethical reasons, data allowing to estimate model parameters are poor, sparse, and do not often concern human populations. Recourse to extrapolation is then needed and physiological toxicokinetic models facilitate the required extrapolations (inter-species, inter-subject, ...). The problem induced by the poor quality of data can be partly circumvented in a Bayesian framework where the available *a priori* values on physiological parameters (blood flows, organ volumes, etc.) can be used.

Keywords : Toxicokinetic models, Bayesian statistics, Stochastic simulations, Toxicology.

* Unité de Toxicologie expérimentale, Institut National de l'Environnement Industriel et des Risques, INERIS, Parc Technologique ALATA, PB2, 60550 Verneuil en Halatte
sandrine.micallef@ineris.fr, frederic.bois@ineris.fr

1. Introduction

Certaines substances chimiques présentes dans l'environnement peuvent engendrer des effets nocifs après absorption. Pour comprendre les mécanismes d'action sous-jacents, des études toxicologiques sont menées afin de recueillir et d'analyser des données expérimentales ou d'observation. Généralement, les contraintes économiques ou éthiques font que les données disponibles pour ce type d'étude sont assez pauvres voire insuffisantes. Il est en effet difficile de mener des expérimentations correspondant à des conditions réelles d'exposition (sur la population humaine par exemple) compte tenu des conséquences nocives que ces expositions peuvent engendrer. De ce fait, divers types d'extrapolations restent nécessaires pour prédire les effets toxiques dans un cadre différent des conditions de l'expérimentation. Globalement, l'extrapolation concerne l'exposition (extrapolation inter-doses, inter-durées, inter-voies d'exposition) ou l'individu (transposition inter-espèces, inter-sujets, intra-individu). Pour cela, les modèles mathématiques restent le meilleur outil. Ils sont utilisés pour tester des hypothèses et en déduire des conséquences pour les populations humaines exposées.

Le mécanisme toxique se décompose en deux étapes : acheminement de la substance dans les différents tissus (toxicocinétique) et action toxique au niveau de certains de ces tissus du corps (toxicodynamie). Les effets toxiques des substances chimiques ne sont pas simplement liés à la concentration d'exposition (dans l'environnement, ou dans l'alimentation) mais dépendent de la quantité de produit atteignant les cellules du corps les plus sensibles à ce produit (Andersen, 1995). Les processus complexes d'absorption, de distribution, de transformation chimique et d'élimination font que cette quantité (ou « dose interne ») n'est pas forcément proportionnelle à la concentration d'exposition (Andersen, 1995; Bailer *et al.*, 1997). La relation entre la dose efficace et la dose d'exposition peut varier en fonction de facteurs tels que l'âge, le sexe ou l'état physiologique (Rowland *et al.*, 1989). Des modèles « toxicocinétiques » (ou « pharmacocinétiques » quand il s'agit de substances thérapeutiques) sont développés pour représenter, comprendre et prévoir la distribution locale et temporelle d'un agent toxique dans un organisme.

Parmi les modèles existants, les modèles physiologiques (Gerlowski *et al.*, 1983; Balant *et al.*, 1990; Andersen, 1995) retiennent particulièrement l'attention des toxicologues. Ces modèles ont été développés pour fournir une dosimétrie au niveau des tissus. Leurs structures et paramètres reposant sur une conceptualisation des mécanismes biologiques, ils offrent une description physiologique de l'organisme. De nombreuses informations (données expérimentales référencées, connaissances physiologiques, etc.) sont disponibles *a priori* dans la littérature scientifique quant aux valeurs de leurs paramètres. L'utilisation de ces informations permet de ne pas multiplier inutilement les expérimentations animales ou humaines. De ce fait, les modèles toxicocinétiques physiologiques (TCP) répondent aux contraintes économiques ou éthiques existant en toxicologie.

Cet article fait le point sur l'utilisation des modèles TCP pour l'analyse de données toxicocinétiques. Dans l'optique d'une démarche d'utilisation de

l'information disponible et d'exploitation des bases de données existantes, l'inclusion d'informations connues *a priori* sera présentée. Dans ce cadre, la statistique bayésienne (Bernardo *et al.*, 1994; Gelman *et al.*, 1995; Bernier *et al.*, 2000) est un outil bien approprié à l'analyse de données toxicocinétiques. La calibration des modèles TCP¹ sera abordée selon le paradigme bayésien. Enfin, une dernière section présente les divers types d'extrapolations menées dans le cadre des études toxicocinétiques.

2. Les données toxicocinétiques

Les études toxicocinétiques produisent des données sur la répartition spatio-temporelle d'une substance toxique dans le corps après administration d'une quantité connue. La collecte d'informations spatiales est, en général, assez limitée (sauf dans le cas de l'analyse d'image). La plupart du temps, la concentration de toxique est mesurée dans un seul tissu ou organe du corps (le sang, par exemple, souvent choisi pour des raisons de commodité) et les quantités rejetées (dans l'urine, l'air exhalé...). Pour un tissu analysé, les mesures sont relevées à différents instants; l'information temporelle est plus abondante que l'information spatiale. La Figure 1 montre un exemple de données toxicocinétiques pour le 1,3-butadiène. Ces données représentent la concentration de 1,3-butadiène dans l'air exhalé par quatre sujets humains suite à une exposition à 5 parties par million (en volume dans l'air, ppm) pendant deux heures. Les mesures ont été relevées jusqu'à une heure au-delà de l'exposition. En plus des mesures de concentrations, diverses covariables telles que la masse corporelle, l'âge ou le sexe sont souvent enregistrées. Enfin, toutes ces données sont observées sur plusieurs individus (animaux ou humains).

3. La modélisation toxicocinétique physiologique

Pour décrire la cinétique d'une substance toxique dans un organisme, différents modèles mathématiques sont disponibles. Tous simplifient la complexité physiologique en subdivisant le corps en compartiments (un compartiment est défini comme une quantité de substance homogène au niveau cinétique). Les modèles toxicocinétiques (TC) se classent en deux groupes : les modèles TC empiriques classiques et les modèles TC physiologiques (TCP).

Les modèles TCP s'appuient sur la connaissance biologique. Le corps est modélisé par un ensemble de compartiments correspondant à des organes (foie, reins, poumons, etc.) ou des tissus (muscles, tissus adipeux, os, etc.). Les fluides biologiques sont les principaux vecteurs des composés toxiques dans le corps. Ainsi, les interconnexions entre compartiments représentent souvent les échanges sanguins entre les différents organes. D'autres échanges (via la lymphe ou par diffusion) sont parfois modélisés. Le toxique passe du sang vers les différents organes plus ou moins facilement selon l'affinité des tissus

1. L'acronyme anglais couramment utilisé pour les modèles TCP est PBTK (Physiologically Based Toxicokinetic).

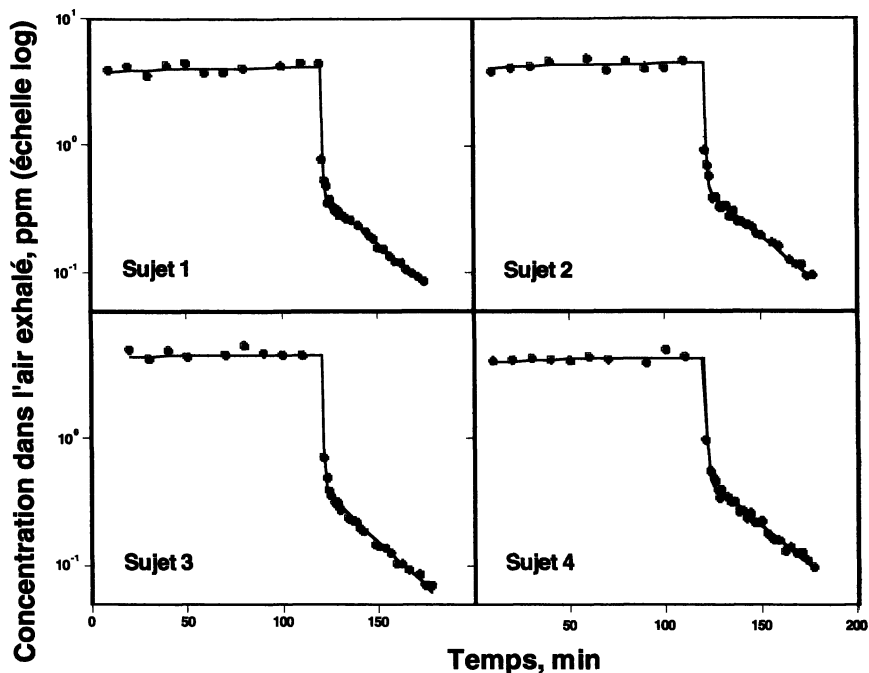


FIG 1. — Concentrations de 1,3-butadiène exhalées suite à une exposition à 5 ppm pendant deux heures. Les données ont été relevées chez quatre volontaires différents et jusqu'à une heure après exposition. Les observations sont représentées par les points; la ligne continue donne les prédictions d'un modèle TCP.

pour la substance. Ce phénomène est caractérisé par un paramètre appelé coefficient «de partage» ou «d'affinité». Les modèles TCP représentent, de plus, les transformations chimiques de la substance par métabolisme. Il s'agit de réactions enzymatiques saturables et potentiellement non linéaires. Le métabolisme a souvent lieu au niveau des cellules du foie, mais peut également se produire dans d'autres organes (reins, poumons, peau, placenta, etc.). Les autres types d'élimination (urinaire, fécale, par expiration ou sudation) sont aussi décrits par les modèles TCP.

Les valeurs des paramètres de ces modèles peuvent être connues à partir de la littérature physiologique ou obtenues par expérimentation. La représentation des organes n'est pas nécessairement exhaustive. Elle dépend de la substance étudiée et des organes sensibles au toxique (organes cibles) ou de ceux jouant un rôle dans le métabolisme ou le stockage de la substance. Leur nombre varie également en fonction de la question toxicologique (hypothèses testées) ou de la connaissance des phénomènes. Dans de nombreux modèles, plusieurs tissus sont agrégés. Cette agrégation est décidée par le modélisateur physiologiste pour des raisons utilitaires ou biologiques (même perfusion sanguine par unité de masse et même affinité pour la substance considérée, par exemple).

Un modèle TCP permet de prédire la concentration de substance dans un tissu donné à partir de la dose d'exposition (quantité ingérée, concentration inhalée, dose injectée par voie intraveineuse, etc.). La fonction paramétrée f_θ reliant la concentration d'exposition, C_{Exp} , et les concentrations internes, C_i , s'écrit symboliquement :

$$C_i = f_\theta(C_{Exp}).$$

Les contraintes du transport des substances sont, en général, modélisées par un système paramétré dont f_θ est la solution. Ce système est composé d'équations différentielles décrivant la variation de concentration du toxique dans chacun des compartiments. Les paramètres θ du système sont :

- les paramètres spécifiques aux caractéristiques physiologiques de l'espèce considérée (par exemple, taux de ventilation pulmonaire noté Fl ou, pour le compartiment i , volume des tissus V_i , flux sanguins F_i)
- et les paramètres spécifiques à la fois au toxique et à l'organisme (coefficient de partition P_i ou constante cinétique de métabolisme k_i , par exemple).

La variation de quantité de toxique dans le compartiment i est décrite, de façon générique, par l'équation différentielle suivante :

$$\frac{\partial Q_i}{\partial t} = F_i \left(C_{e_i} - \frac{Q_i}{V_i P_i} \right) + g_i(t) - h_{k_i}(t) - l_i(t)$$

où C_{e_i} , est la concentration entrant par l'intermédiaire du sang, la fonction g_i caractérise l'administration, h_{k_i} la transformation par métabolisme dont le paramètre est noté k_i et l_i correspond à l'élimination (non nulles lorsqu'elles ont lieu dans le compartiment i). Ce système, résolu en général par intégration numérique avec des valeurs de paramètres, permet de simuler les concentrations internes au cours du temps dans tous les compartiments du modèle. La Figure 2 présente un modèle physiologique à 7 compartiments décrivant l'exposition par inhalation à de nombreux agents gazeux. Les échanges avec le milieu extérieur se font ici via les poumons (respiration), les reins (excrétion urinaire) et le foie (métabolisme hépatique). La plupart des compartiments sont interconnectés parallèlement. Les équations différentielles caractérisant les échanges pour chaque compartiment sont indiquées directement sur la Figure 2. La résolution numérique de ces équations fournit, dans chaque compartiment, la concentration ou la quantité de toxique en fonction du temps. Des courbes de cinétique simulée sont proposées sur la Figure 1, en trait continu. Elles donnent la concentration de 1,3-butadiène dans l'air exhalé en fonction du temps suite à une exposition à 5 ppm pendant deux heures.

Les modèles TCP peuvent, en outre, décrire des phénomènes physiologiques tels que les phénomènes de premier passage ou les cycles entéro-hépatiques. Après l'absorption, une certaine quantité de substance toxique se propage dans le corps. Une partie de cette quantité atteint les sites d'élimination avant passage par les autres tissus. Cette « perte » correspond au phénomène de premier passage et doit être prise en compte lorsque l'on tente des extrapolations entre voie d'administration, en particulier. Pour cela, le modèle TCP doit comporter une description, même simple, des compartiments concernés.

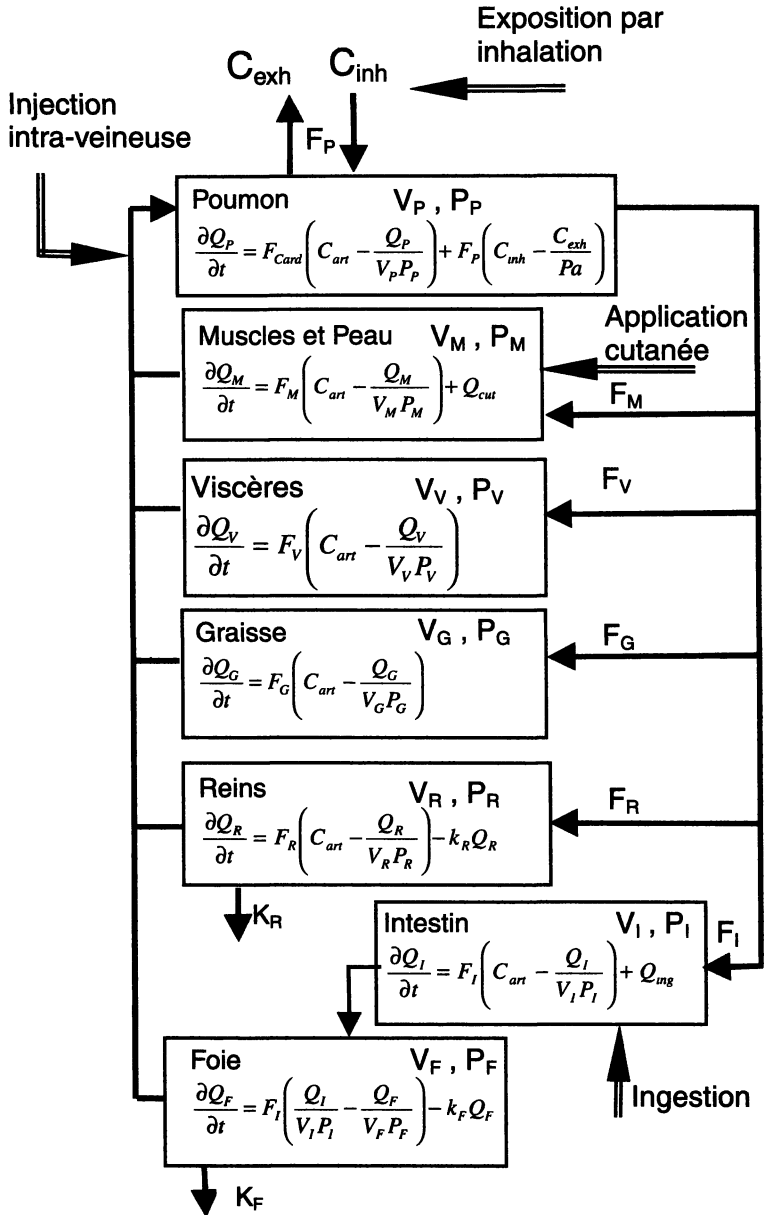


FIG 2. — Exemple d'un modèle TCP générique à 7 compartiments. Quatre voies d'exposition sont représentées : intra-veineuse, cutanée, par ingestion et par inhalation. Ce modèle permet de décrire des expositions simultanées si nécessaire et peut être utilisé dans le cadre d'une transposition inter-voies d'exposition. Les paramètres de ce modèle sont les volumes des compartiments notés V, les coefficients de partages P, les flux sanguins F. Les quantités de chaque compartiment ou les quantités administrées sont représentées par la lettre Q. Deux types de métabolisme sont représentés : le métabolisme rénal caractérisé par la constante de réaction K_R et le métabolisme hépatique par K_F .

La description des mécanismes de distribution de l'agent toxique est plus ou moins fine selon les modèles. Elle varie principalement selon les objectifs de l'étude (connaître une quantité métabolisée, fournir des prédictions de concentrations au niveau de tel organe...). Pour un même agent, différentes structures de modèles TCP sont donc admissibles. Le 1,3-butadiène, par exemple, est un agent pour lequel différents modèles TCP ont été définis. Mezzetti *et al.* (2003) utilise un modèle TCP comprenant trois compartiments dans lequel le compartiment des organes « bien perfusés » est le seul site de métabolisme. D'autres auteurs (Kohn *et al.*, 1993) ont proposé un modèle à six compartiments différenciant les métabolismes hépatique, viscéral et pulmonaire. Dans un travail plus récent, Kohn *et al.* (1996) discutent de l'effet de la structure sur les prédictions des modèles TCP pour le 1,3-butadiène. Malgré leurs différentes structures, les modèles précédents ont permis de répondre à des questions toxicologiques précises concernant la même substance.

4. Modélisation statistique et modélisation TCP

4.1. L'information *a priori*

De par leur structure paramétrique, les modèles TCP permettent de prendre en compte un savoir physiologique *a priori*. L'information *a priori* regroupe les connaissances et l'expertise relatives à l'étude et recueillies avant l'expérimentation. Il s'agit généralement d'information concernant le mécanisme toxique (passage d'une barrière physiologique, transformation chimique des xénobiotiques dans les cellules du foie, par exemple) ou les grandeurs physiologiques (masse du foie, volume des reins, etc.). Ces informations peuvent être prises en compte au niveau de la structure du modèle ou de la valeur de ses paramètres. Qu'elles portent sur la valeur des paramètres anatomo-physiologiques ou sur le mécanisme lui-même, toutes ces informations comprennent une part d'incertitude qui doit être intégrée au modèle. Le recours à une simple estimation *a priori* des incertitudes sur les paramètres du modèle (Bois, 2003) permet de donner des intervalles de confiance pour les prédictions en utilisant des méthodes de Monte Carlo (Hammersley and Handscomb, 1964).

La littérature scientifique offre une source considérable d'information *a priori* utilisable pour la construction des modèles TCP. De nombreuses publications (Arms *et al.*, 1988; Davies *et al.*, 1993; Brown *et al.*, 1997; International Commission on Radiological Protection (ICRP), 2003) recensent des valeurs physiologiques (volumes d'organes, flux sanguins, etc.) pour l'humain et les animaux communément utilisés dans les laboratoires. La plupart des publications donnent des valeurs de référence et peuvent aller jusqu'à indiquer des intervalles de valeurs biologiques plausibles (Brown *et al.*, 1997). Ces intervalles fournissent une mesure de l'incertitude associée au paramètre considéré. Il est cependant exceptionnel que ces documents informent sur la variabilité des paramètres dans la population. Cela est regrettable d'autant que les études toxicologiques n'observent qu'un échantillon restreint de la population qu'elles visent à protéger.

4.2. Utilisation de l'information provenant d'expérimentations pour l'estimation des paramètres

Les valeurs des paramètres des modèles TCP peuvent aussi être déterminées expérimentalement (par exemple, pesée des sujets pour évaluer la masse corporelle). Malgré leur base physiologique, ces valeurs ne sont pas toujours observables directement (constante de vitesse d'élimination ou masse d'un organe interne, par exemple). Dans ce cas, l'information apportée par des données expérimentales sur le comportement de variables d'état du modèle (*e.g.* concentrations sanguines en fonction du temps) devient nécessaire même si ces observations ne concernent pas directement les valeurs des paramètres inconnus. L'estimation des paramètres consiste alors à extraire l'information de ces données expérimentales et peut être menée dans un cadre fréquentiste (par « ajustement ») ou bayésien (par « calibration »).

Certains modélisateurs refusent d'ajuster les modèles TCP car leur qualité pourrait être détériorée par de mauvaises données (Hopkins *et al.*, 1996). Ils prétextent que l'ajustement du modèle pourrait transférer les erreurs (ou imprécisions) contenues dans les données vers les paramètres et fausser ses prédictions. D'autres arguments en défaveur de l'ajustement concernent les erreurs de modélisation elles-mêmes. Premièrement, si un modèle TCP est construit sur la base d'hypothèses mécanistes incorrectes, son ajustement pourrait dissimuler les erreurs commises à ce niveau. Deuxièmement, un modèle TCP ajusté sur la base de données récoltées dans certaines conditions expérimentales, peut donner des prédictions différentes de nouvelles données produites dans les mêmes conditions (au moins en apparence!).

Ces remarques ne prennent pas en compte la possibilité de compléter la modélisation physiologique par une modélisation statistique (modèle d'erreur, prise en compte des variabilités explicables, etc.). Malgré tout, ajuster les modèles TCP reste discutable. En effet, ces modèles sont complexes, en général sur-paramétrés par rapport aux jeux de données disponibles et les méthodes d'ajustement classiques sont alors assez inefficaces. Elles tendent à redécouvrir des mécanismes bien connus (effet du poids, par exemple) en utilisant des données coûteuses. Enfin, l'ajustement n'est pas nécessaire pour évaluer la précision des prédictions, puisque des simulations stochastiques simples le permettent (Hammersley and Handscomb, 1964; Greenland, 2001; Sweeney *et al.*, 2001; Bois, 2003).

4.3. Estimation : l'approche bayésienne

La nécessité d'estimer les paramètres du modèle, le souhait de prendre en compte l'information disponible *a priori* et de tenir compte des incertitudes conduisent à l'utilisation de la statistique bayésienne pour l'analyse des données toxicocinétiques (Bois *et al.*, 1996; Gelman *et al.*, 1996).

Globalement, l'analyse bayésienne (Bernardo and Smith, 1994; Gelman *et al.*, 1995; Gilks *et al.*, 1996; Bernier *et al.*, 2000) permet de construire une inférence à partir de données sur la base d'un modèle de probabilité pour des quantités observables ou non-observables. Sa principale caractéristique est la

quantification des incertitudes par des densités de probabilité. Le processus bayésien d'analyse des données inclut plusieurs étapes. La première consiste à définir un modèle de probabilité pour toutes les quantités (observables ou non) du modèle. C'est au cours de cette étape que l'expertise ou les informations *a priori* provenant de la littérature (ou des bases de données) peuvent être incluses en les traduisant en terme de densités de probabilité. L'ensemble des densités définies pour chaque paramètre du problème constitue la *densité a priori* qui est plus ou moins dispersée en fonction de la précision des connaissances sur les paramètres considérés. La seconde étape consiste à conditionner les paramètres sur les données observées par l'intermédiaire d'un calcul de vraisemblance. À partir du théorème de Bayes, la densité jointe *a priori* est mise à jour en utilisant la vraisemblance des données pour fournir une densité jointe *a posteriori*. Cette densité jointe *a posteriori* est donc une distribution de probabilité conditionnelle des quantités inobservables d'intérêt sachant les données collectées. La détermination de cette loi *a posteriori* constitue l'estimation bayésienne d'un modèle ou « calibration ». De la densité *a posteriori*, on peut dériver divers estimateurs des paramètres et les incertitudes qui leur sont associées. La troisième étape de la démarche bayésienne repose sur la critique du modèle calibré et de ses prédictions. Plusieurs questions peuvent être posées dans ce cadre : Le modèle s'ajuste-t-il correctement aux observations ? Ses prédictions sont-elles raisonnables ? Les conclusions qui en découlent sont-elles cohérentes avec la connaissance du problème ? Et enfin, les résultats obtenus sont-ils sensibles aux hypothèses faites pour la construction du modèle ?

L'approche bayésienne s'applique particulièrement bien dans le contexte paramétrique des modèles TCP. Tout d'abord, elle permet d'inclure la connaissance *a priori* sur les paramètres dans le cadre de l'analyse de nouvelles données expérimentales et, de ce fait, utilise au mieux l'information apportée par des expériences souvent coûteuses. Ensuite, elle conduit naturellement à des estimations de paramètres ou des prédictions avec un intervalle de crédibilité associé. La calibration bayésienne diminue les biais du modèle qui pourraient exister sur la base de simulations reposant uniquement sur les paramètres *a priori* et permet de discuter ses prédictions. Par ailleurs, tous les paramètres d'un modèle ne bénéficiant pas nécessairement d'une information *a priori*, certains nécessitent d'être estimés. Enfin, elle est maintenant facilement réalisable par des méthodes numériques (Bois *et al.*, 1997) fondées sur des algorithmes de simulation de Monte Carlo par chaîne de Markov (Gilks *et al.*, 1996 ; Robert, 1996) et peut être conduite malgré la complexité des modèles. Cette approche est donc particulièrement recommandée pour l'estimation des modèles TCP même s'ils tendent à devenir de plus en plus complexes.

5. Application de modèles TCP à la prédiction

Les modèles TCP ont souvent pour objectif d'extrapoler des conditions expérimentales vers les situations environnementales. L'analyse de données obtenues après expositions expérimentales et/ou la détermination d'un mécanisme d'action opérant chez l'animal ne serait guère utile en l'absence de procédures

scientifiques d'extrapolation (ou « transposition ») des résultats à l'homme. Il s'agit alors de statuer sur des expositions plus faibles, plus fortes ou sur des expositions par inhalation quand les données concernent l'ingestion. Même lorsque des données humaines sont disponibles, la question se pose par exemple de leur validité pour la femme, la femme enceinte, l'enfant, et autres populations sensibles (Louis, 1991 ; Oflaherty, 1997). La recherche en toxicologie se concentre sur l'étude des mécanismes d'action et tend à considérer qu'un modèle mécaniste conduit à des extrapolations plus robustes qu'un modèle empirique. L'incorporation de considérations statistiques sur la variabilité des valeurs de paramètres physiologiques dans les populations considérées (Sheiner *et al.*, 1982 ; Gelman *et al.*, 1996 ; Wakefield *et al.*, 1996 ; Racine-Poon *et al.*, 1998) peut également intervenir dans le processus d'extrapolation.

Dans la suite, les différents types d'extrapolation sont détaillés, en particulier les extrapolations inter-doses, inter-durées d'exposition, inter-voies d'administration, inter-espèces, inter-sujets et intra-individuelle.

5.1. Extrapolations inter-doses et inter-durées

La dose interne est la principale variable décrite par les modèles toxicocinétiques (qu'ils soient physiologiques ou non) puisqu'elle conditionne l'effet toxique. Elle est à son tour conditionnée par le niveau de concentration externe qui est, lui-même, une fonction arbitraire du temps. Par exemple, l'exposition peut être unique ou répétée, se faire à concentration fixe ou fluctuante. Les doses internes atteintes lors d'expérimentations sont souvent très différentes de celles rencontrées dans les situations réelles d'exposition sur les lieux de travail ou dans l'environnement. Pour extrapoler d'une exposition donnée à une autre prédite, il « suffit » de changer, en entrée du modèle, le niveau de concentration d'exposition, la durée d'exposition, sa fréquence, etc.

Néanmoins, une extrapolation entre faible et forte dose reste bien sûr délicate car l'on s'intéresse à ce qui se passe en dehors de l'intervalle d'observation du phénomène. Les modèles empiriques ajustés peuvent fournir une interpolation sur la plage d'observation des données. Hors de cet intervalle, tout modélisateur est conscient qu'il peut, hélas, s'éloigner considérablement du phénomène réel qu'il cherche à prédire. Les modèles physiologiques sont, en revanche, théoriquement capables de donner des prédictions correctes même en dehors de l'intervalle de collection des données puisqu'ils s'appuient sur des mécanismes physiologiques supposés être indépendants de l'exposition.

5.2. Transpositions inter-voies d'exposition

Les expérimentations dans lesquelles la voie d'exposition n'est pas la voie d'intérêt peuvent également être mises à profit grâce à la modélisation TCP. Une fois dans le corps, la substance se distribue selon les mêmes mécanismes indépendamment de la voie d'administration. Le système d'équations différentielles (et ses paramètres associés) caractérisant la distribution de la substance est alors le même quelle que soit la voie d'exposition considérée. On peut donc prédire la toxicocinétique d'un agent inhalé sur la base d'études utilisant la voie d'exposition cutanée. Il suffit pour cela que le modèle décrive les voies

concernées et les flux de substance correspondants. Le modèle physiologique de la Figure 2 est ainsi capable de simuler des expositions par inhalation, application cutanée, ingestion et intra-veineuse. L'exposition par inhalation est, par exemple, représentée par les échanges entre le poumon et l'air ambiant. Il est même possible de prédire des expositions simultanées par ces différentes voies.

5.3. Transposition inter-espèces

Pour des raisons éthiques évidentes et pour des raisons techniques, il est difficile voire impossible d'exposer des volontaires humains et d'observer des doses internes. C'est pourquoi la plupart des études toxicologiques visant à déterminer des doses admissibles (doses maximales ne causant pas d'effet toxique) utilisent des systèmes *in vitro* ou ont recours à l'expérimentation animale. Les études *in vitro* permettent de comprendre le mécanisme toxique local (au niveau des cellules cibles, par exemple). Cela n'est pas toujours suffisant car il reste à comprendre le phénomène de distribution jusqu'à ces tissus cibles. C'est alors que l'expérimentation animale devient utile. Malgré tout, des différences significatives (qualitatives et quantitatives) existent entre les physiologies humaine et animale. Pour transposer les doses admissibles obtenues d'une espèce à une autre, deux solutions sont envisageables : l'utilisation de « facteurs d'incertitude » ou la modélisation physiologique. La première solution consiste à diviser par un facteur (10, en général) les doses obtenues chez l'animal pour les transposer à l'homme. Ce facteur, choisi plus ou moins arbitrairement, est souvent discutable. En revanche, la modélisation TCP permet, par sa description paramétrique de l'organisme, d'expliquer les différences inter-espèces. Le principe de base de transposition par modélisation TCP (Lawrence *et al.*, 1997; Wang *et al.*, 2000) est de supposer que les mécanismes (de pénétration, distribution, métabolisme et excrétion) sont identiques chez les espèces considérées et que les différences concernent uniquement les grandeurs physiologiques. Des différences de mécanisme connues (par exemple, de métabolisme) peuvent également être prises en compte en adaptant le modèle. Les procédures de transposition consistent donc à conserver ou adapter la structure du modèle et à modifier la valeur de ses paramètres physiologiques (volumes des organes, flux plasmatiques...) en fonction de l'espèce (Ings, 1990; Watanabe *et al.*, 1996) (par l'intermédiaire de bases de données existantes, par exemple : Brown *et al.*, 1997).

Néanmoins, la valeur de certains paramètres n'est pas toujours connue et reste difficile à transposer d'une espèce à l'autre. C'est le cas, par exemple, du facteur d'irrigation de certains tissus chez le fœtus en développement. Dans ce contexte, le recours à l'extrapolation « allométrique » est encore possible (Davidson *et al.*, 1986). L'extrapolation allométrique considère que la valeur du paramètre d'intérêt s'écrit comme une fonction paramétrée de la masse corporelle :

$$P = a \cdot W^k$$

où P est la valeur extrapolée du paramètre considéré, a et k des constantes indépendantes de l'espèce et W la masse corporelle de l'espèce. Les équations

allométriques sont généralement paramétrées par ajustement des paramètres a et k sur la base de données expérimentales ou provenant de la littérature.

5.4. Extrapolation inter-sujets ou intra-individuelle

Malgré les difficultés éthiques, certaines études expérimentales de toxicologie observent directement des populations humaines. Ces études expérimentales sont très contrôlées et un comité d'éthique doit toujours être consulté. Dans ce type d'étude, la dose d'exposition est comparable ou inférieure aux doses environnementales habituelles. Néanmoins, l'accroissement possible du risque individuel pour les volontaires ainsi que le coût du montage de l'expérience conduisent à limiter leur recrutement. L'information recueillie suite à l'exposition des sujets est intéressante mais c'est plutôt la population dans sa globalité que l'on souhaite protéger et non les volontaires eux-mêmes (en tant qu'individus ayant participé à l'expérience, sauf dans un cadre thérapeutique). L'analyse sur un échantillon de population doit donc être utile pour renseigner sur l'ensemble de la population malgré les variations inter-individuelles.

On peut aussi s'intéresser à la toxicocinétique d'un agent chez des populations notoirement différentes de celles observées. Ces différences peuvent concerner le sexe, l'âge, l'ethnie ou l'état physiologique des sujets. Par exemple, les femmes enceintes ne sont jamais volontairement exposées, mais on peut vouloir inférer la toxicocinétique d'une substance chez elles. On peut s'intéresser, également, à la toxicité du plomb chez les enfants et ne disposer que de données chez l'adulte. Les différences expliquées entre les individus sont prises en compte par l'introduction de covariables dans le modèle. Ces covariables permettent également de traiter les changements intra-individuels observables entre deux expérimentations (variation de la masse, de l'âge, de l'état physiologique d'un même sujet). Lorsque l'utilisation des covariables ne permet pas d'extrapoler, la modélisation déterministe des paramètres peut être utilisée pour les extrapolations inter-sujets et intra-individuelle. Cependant, les covariables ou les paramètres ne suffisent pas toujours à expliquer les différences dans la réponse toxicocinétique entre deux sujets ou un même individu. Pour les fluctuations inexpliquées, il est nécessaire de prendre en compte les variabilités inter et/ou intra-individu par modélisation statistique.

Extrapolation par modélisation déterministe

La modélisation physiologique peut être étendue pour décrire les différences inter-sujets affectant les valeurs de paramètres. Pour une extrapolation entre deux sujets d'âge différent, par exemple, une prise en compte de l'évolution de la masse des différents tissus ou organes en fonction de l'âge permet d'adapter automatiquement la valeur des paramètres du modèle physiologique. L'évolution de la masse des tissus chez l'homme et la femme a ainsi pu être utilisée pour prédire l'accumulation de 2,3,7,8-tetrachloro-*p*-dioxine (TCDD) à long terme (Van der Mollen *et al.*, 1996). Dans le même ordre d'idée, nous avons développé un modèle physiologique (Figure 3) pour décrire les changements biologiques au cours de la gestation, basé sur un modèle de Luecke *et al.*

(1997). Les compartiments grisés (correspondant au fœtus et à certains tissus maternels) sont ceux pour lesquels un accroissement a été pris en compte à partir d'une modélisation déterministe. Ce modèle permet une simulation de la toxicocinétique de substances diverses chez la femme enceinte et son fœtus.

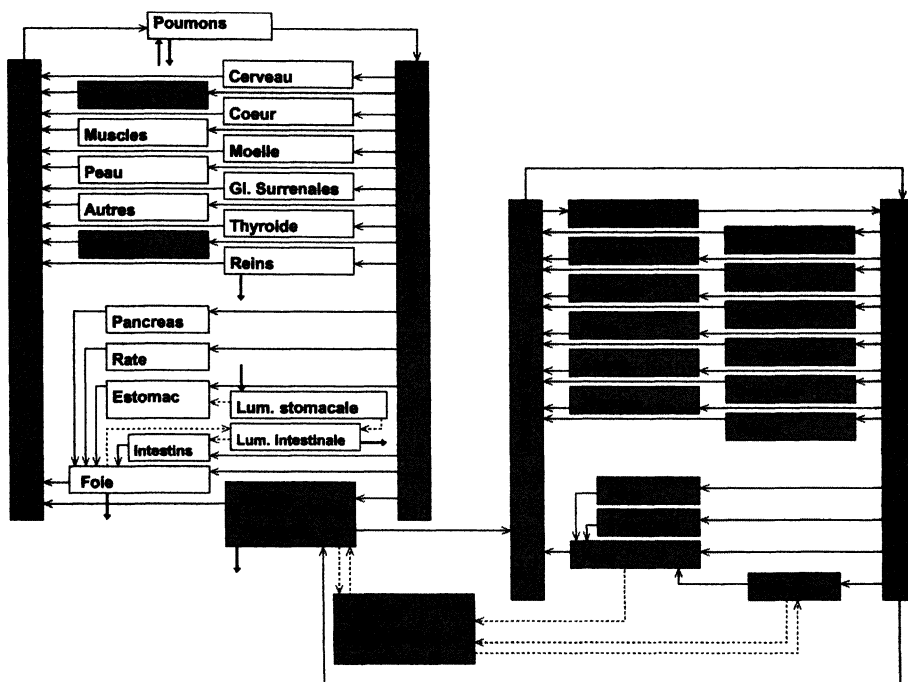


FIG 3. — Modèle physiologique décrivant la gestation. Le sous-modèle gauche représente la mère et le sous-modèle droit son fœtus. Les compartiments grisés correspondent aux tissus pour lesquels la croissance au cours de la gestation est modélisée de façon déterministe. Les flux de matière sont représentés par des flèches : les traits continus représentent les échanges sanguins et les flèches discontinues les échanges non sanguins (diffusion, passage mécanique, etc.).

Extrapolation par modélisation statistique hiérarchique

La modélisation déterministe n'est applicable qu'à ce que l'on connaît des causes de l'hétérogénéité. Or, par manque d'information, la variabilité n'est pas toujours interprétable d'un point de vue biologique. Dans ce cas, la modélisation statistique et le calcul des probabilités permettent la prise en compte de l'hétérogénéité à l'aide d'outils tels que les modèles de population (Sheiner, 1984; Racine-Poon *et al.*, 1990; Aarons, 1992). Ces modèles correspondent aux modèles à effets mixtes en statistique classique ou aux modèles hiérarchiques en statistique bayésienne. Ils ont pour objectif d'extraire une description quantitative de la variabilité dans une large population, à partir de données individuelles. Le principe de base est de considérer que les valeurs d'un (ou des) paramètre(s) sont distribuées aléatoirement dans la population. Chaque paramètre individuel est situé autour de la valeur moyenne de population. D'un

individu à l'autre, les paramètres diffèrent et caractérisent le degré de ressemblance entre les individus. Cette variabilité inter-individus peut être décrite par une loi de probabilité multivariée appelée « distribution de population ». La transposition par modélisation statistique se traduit généralement par la proposition de modèles multi-niveaux de variabilité inter-sujets. Un modèle typique de population est présenté sur la Figure 4. Deux niveaux de hiérarchie sont différenciés : le niveau individuel (sujets) et le niveau de population. Au niveau individuel, les données Y sont mesurées aux temps t , après une exposition E . Certaines caractéristiques du sujet peuvent être connues précisément (comme, par exemple, l'âge ou la masse corporelle), auquel cas elles sont considérées comme des covariables (Φ). Si elles sont inconnues, elles deviennent des paramètres et sont notées θ . Les caractéristiques individuelles (paramètres θ ou covariables Φ) conditionnent les observations. Au niveau population, la variabilité inter-sujets est décrite en considérant que les paramètres θ , eux-mêmes, sont les réalisations d'un vecteur de variables aléatoires issues de distributions de population dont les paramètres (μ et Σ) règlent la plus ou moins grande ressemblance entre individus.

Dans cette représentation, deux sous-modèles sont définis :

- Le modèle statistique lie les aspects prévisibles des données aux observations mesurées. Les différences entre prévisions et données sont supposées provenir de l'incertitude sur les mesures des erreurs de modélisation ou des variabilités intra-individuelles. Ces déviations peuvent, par exemple, être modélisées par des variables aléatoires normales de moyenne nulle et de variance σ^2 . L'approche bayésienne permet de prendre en compte les incertitudes sur les valeurs de paramètres (de population par exemple, μ et Σ ou d'autres sur σ^2) en utilisant des densités de probabilité *a priori* notées P .
- Le modèle TCP, f , lie les variables prédites (concentrations ou quantités) aux variables E , t , θ et Φ . f correspond à la solution du système d'équations différentielles défini à partir de la modélisation toxicocinétique. Dans le cas où les modèles toxicocinétiques sont assez simples (modèle à un compartiment, par exemple), il est possible de résoudre le système d'équations différentielles analytiquement. Si ce n'est pas le cas, des méthodes de résolution numérique sont souvent utilisées.

Lors d'études toxicologiques antérieures, nous avons utilisé les modèles hiérarchiques de populations incluant la modélisation TCP et traité l'estimation des paramètres de ces modèles dans un cadre bayésien numérique (Bois *et al.*, 1996; Bois *et al.*, 1996; Gelman *et al.*, 1996).

6. Conclusion et perspectives

La modélisation TCP cherche à représenter un phénomène physique en fonction de différentes covariables physiologiques (masse corporelle, volume des tissus, etc.). Un large corpus de données informant sur la valeur moyenne de ces covariables dans la population a été accumulé au cours du siècle dernier,

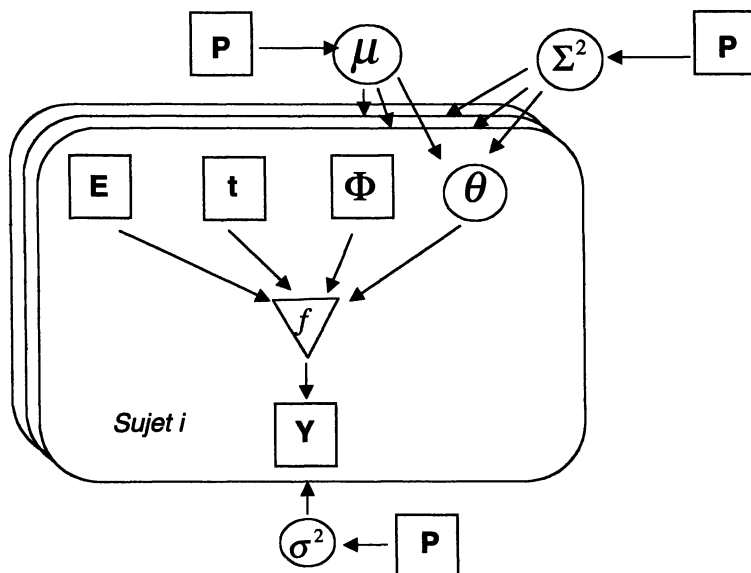


FIG 4. — Modèle multi-niveaux de variabilité inter-sujets. Au niveau individuel, les données (Y) sont mesurées aux temps t , après une exposition E . Les caractéristiques individuelles sont notées θ ou Φ et conditionnent les observations. f représente le modèle toxicocinétique. Au niveau population, la variabilité inter-sujets est décrite en considérant que les paramètres θ sont les réalisations d'un vecteur de variables aléatoires issues de distributions statistiques dont les paramètres de population sont μ et Σ . Le modèle d'erreur est caractérisé par la variabilité σ^2 . Aux paramètres μ , Σ et σ^2 sont associées des distributions *a priori*, notées P . Trois types de nœuds sont représentés sur ce graphe : les nœuds circulaires représentent les variables inconnues, les carrés correspondent à ce qui est supposé connu sans incertitude et le triangle représente le modèle déterministe. Les flèches correspondent à des liens déterministes ou statistiques.

mais ces informations restent éparées. Le développement d'Internet devrait permettre une meilleure accessibilité de ces données antérieures. En parallèle, la modélisation physiologique offre la possibilité de mieux structurer et utiliser cette information.

La recherche en toxico/pharmaco-cinétique s'oriente donc vers des descriptions de plus en plus fines de l'organisme, capables d'intégrer davantage de données (comme, par exemple, le modèle de la femme enceinte présenté Figure 3). La création de modèles toxicocinétiques à la fois génériques et détaillés permettrait de décrire, à l'aide d'un même modèle, la distribution de multiples substances. En outre, l'imagerie médicale fournit maintenant des données sur la distribution de toxiques au niveau tissulaire et sur l'anatomie ou la physiologie de sujets, sans dommage pour leur santé. De telles données vont permettre à leur tour de calibrer des modèles plus complexes.

Néanmoins, les informations sur la distribution des paramètres physiologiques dans la population sont encore assez pauvres. Par exemple, la forme précise des lois de distribution est la plupart du temps inconnue. Les modèles

de population généralement utilisés pour l'analyse des données toxicocinétiques supposent connues les lois de distribution des paramètres de population. En général, des lois normales, log-normales ou Gamma sont utilisées. Ces hypothèses mériteraient d'être relaxées par l'utilisation de modèles de population non paramétriques (Claret *et al.*, 1996; Wakefield *et al.*, 1997; Racine-Poon and Wakefield, 1998).

L'analyse toxicocinétique de population ou l'optimisation de plans d'expériences sont des outils utiles aux études de toxicologie expérimentale (Bois *et al.*, 1999). Elles requièrent la mise en place de modèles statistiques dont les paramètres ont besoin d'être estimés. L'utilisation de l'information *a priori* pour l'estimation des paramètres donne tout son sens à la calibration bayésienne (Gelman *et al.*, 1996; Bois *et al.*, 1999). La possibilité de calibrer les modèles TCP permet aussi de les critiquer et/ou de les comparer entre eux. Bien que la définition du processus de « validation » d'un modèle reste une question ouverte, il est clair que ces étapes lui sont essentielles et le sont, par suite, pour l'utilisation de ses prédictions.

L'objectif des études de toxicologie est de déterminer l'incidence des expositions sur la santé des populations. Leur finalité est donc, plus précisément, de quantifier les effets des toxiques. Les modèles TCP, élaborés pour déterminer les doses internes (en particulier, au niveau des cellules cibles de l'action toxique), ne décrivent qu'une partie du processus et ne permettent pas de lier l'exposition à l'effet. Il faut faire appel à des modèles dose-réponse (ou « toxicodynamiques ») pour établir une relation entre les doses internes et les effets qu'elles induisent. Le couplage de modèles TCP et de modèles dose-réponse est un axe de recherche actuel en toxicologie et en pharmacologie.

7. Références

- AARONS L. (1992). Population pharmacokinetics. *International Journal of Clinical Pharmacology, Therapy, and Toxicology* **30**, 520-522.
- ANDERSEN M. E. (1995). Development of physiologically based pharmacokinetic and physiologically based pharmacodynamic models for applications in toxicology and risk assessment. *Toxicology Letters* **79**, 35-44.
- ARMS A. D. and C. C. TRAVIS (1988). *Reference Physiological Parameters in Pharmacokinetics modelling*.
- BAILER A. J. and D. A. DANKOVIC (1997). An introduction to the use of physiologically based pharmacokinetic models in risk assessment. *Statistical Methods in Medical Research* **6**, 341-358.
- BALANT L. P. and M. GEX-FABRY (1990). Physiological pharmacokinetic modelling. *Xenobiotica* **20**, 1241-1257.
- BERNARDO J. M. and A. F. M. SMITH (1994). *Bayesian Theory*. New York, Wiley.
- BERNIER J., E. PARENT and J. J. BOREUX (2000). *Traitement bayésien des incertitudes*. Paris, Lavoisier.
- BOIS F. Y. (2003). Modélisation toxicocinétique de la concentration de 2,3,7,8-tetrachloro-p-dioxine après ingestion chez la femme. *Environnement, Risques et Santé* **2**(1), 45-53.

- BOIS F. Y., A. GELMAN, J. JIANG, *et al.* (1996). Population toxicokinetics of tetrachloroethylene. *Archives of Toxicology* **70**(6), 347-355.
- BOIS F. Y., E. JACKSON, K. PEKARI, *et al.* (1996). Population toxicokinetics of benzene. *Environmental Health Perspectives* **104** (suppl. 6), 1405-1411.
- BOIS F. Y. and D. MASZLE (1997). MCSim : a simulation program. *Journal of Statistical Software* **2**(9),
http://toxi.ineris.fr/activites/toxicologie_quantitative/mcsim/mcsim.php.
- BOIS F. Y., T. J. SMITH, A. GELMAN, *et al.* (1999). Optimal design for a study of butadiene toxicokinetics in humans. *Toxicological Sciences : an Official Journal of the Society of Toxicology* **49**(2), 213-224.
- BROWN R. P., M. D. DELP, S. L. LINDSTEDT, *et al.* (1997). Physiological parameter values for physiologically based pharmacokinetic models. *Toxicology and Industrial Health* **13**(4), 407-484.
- CLARET L. and A. ILIADIS (1996). Nonparametric density estimation applied to population pharmacokinetics. *Math. Biosci.* **133**(1), 51-68.
- DAVIDSON I. W. F., J. C. PARKER and R. P. BELILES (1986). Biological basis for extrapolation across mammalian species. *Regulatory Toxicology and Pharmacology* **6**, 211-237.
- DAVIES B. and T. MORRIS (1993). Physiological parameters in laboratory animals and humans. *Pharmaceutical Research* **10**, 1093-1095.
- GELMAN A., F. Y. BOIS and J. JIANG (1996). Physiological pharmacokinetic analysis using population modeling and informative prior distributions. *Journal of the American Statistical Association* **91**, 1400-1412.
- GELMAN A., J. B. CARLIN, H. S. STERN, *et al.* (1995). *Bayesian Data Analysis*. London, Chapman & Hall.
- GERLOWSKI L. E. and R. K. JAIN (1983). Physiologically based pharmacokinetic modelling : principles and applications. *Journal of Pharmaceutical Sciences* **72**, 1103-1127.
- GILKS W. R., S. RICHARDSON and D. J. SPIEGELHALTER (1996). *Markov Chain Monte Carlo in Practice*. London, Chapman & Hall.
- GREENLAND S. (2001). Sensitivity analysis, Monte Carlo risk analysis, and Bayesian uncertainty assessment. *Risk Analysis : an Official Publication of the Society for Risk Analysis* **21**(4), 579-583.
- HAMMERSLEY J. M. and D. C. HANDSCOMB (1964). *Monte Carlo Methods*. London, Chapman and Hall.
- HOPKINS J. C. and R. J. LEIPOLD (1996). On the dangers of adjusting the parameter values of mechanism-based mathematical models. *Journal of Theoretical Biology* **183**, 417-427.
- INGS R. M. (1990). Interspecies scaling and comparisons in drug development and toxicokinetics. *Xenobiotica* **20**, 1201-1231.
- International Commission on Radiological Protection (ICRP) (2003). *Basic Anatomical and Physiological Data for Use in Radiological Protection : Reference Values*. Stockholm, J. Valentin.
- KOHN M. C. and R. L. MELNICK (1993). Species differences in the production and clearance of 1,3-butadiene metabolites : a mechanistic model indicates predominantly physiological, not biochemical, control. *Carcinogenesis* **14**(4), 619-628.

- KOHN M. C. and R. L. MELNICK (1996). Effects of the structure of a toxicokinetic model of butadiene inhalation exposure on computed production of carcinogenic intermediates. *Toxicology* **113**(1-3), 31-39.
- LAWRENCE G. S. and F. A. GOBAS (1997). A pharmacokinetic analysis of interspecies extrapolation in dioxin risk assessment. *Chemosphere* **35**(3), 427-452.
- LOUIS T. A. (1991). Assessing, accommodating, and interpreting the influences of heterogeneity. *Environmental Health Perspectives* **90**, 215-222.
- LUECKE R. H., W. D. WOSILAIT, B. A. PEARCE, *et al.* (1997). A computer model and program for xenobiotic disposition during pregnancy. *Computer Methods and Programs in Biomedicine* **53**(3), 201-224.
- MEZZETTI M., J. G. IBRAHIM, F. Y. BOIS, *et al.* (2003). A Bayesian compartmental model for the evaluation of 1,3-butadiene metabolism. *Journal of the Royal Statistical Society, Series C-Applied Statistics* **52**, 291-305.
- OFLAHERTY E. J. (1997). Pharmacokinetics, pharmacodynamics, and prediction of developmental abnormalities. *Reproductive Toxicology* **11**(2-3) : 413-416.
- RACINE-POON A. and A. F. SMITH (1990). *Population models. Statistical Methodology in the Pharmaceutical Sciences*. D. A. Berry, New York, Marcel Dekker, Inc., 139-162.
- RACINE-POON, A. and J. WAKEFIELD (1998). Statistical methods for population pharmacokinetic modelling. *Statist. Methods Med. Res.* **7**(1), 63-84.
- ROBERT C. (1996). *Méthodes de Monte Carlo par Chaînes de Markov*. Paris, Economica.
- ROWLAND M. and T. N. TOZER (1989). *Clinical Pharmacokinetics : Concepts and Applications*. Philadelphia, Lea & Febiger.
- SHEINER L. B. (1984). The population approach to pharmacokinetic data analysis : rationale and standard data analysis methods. *Drug Metabolism Reviews* **15**, 153-171.
- SHEINER L. B. and S. L. BEAL (1982). Bayesian individualization of pharmacokinetics : simple implementation and comparison with non-Bayesian methods. *Journal of Pharmaceutical Sciences* **71**, 1344-1348.
- SWEENEY L. M., T. R. TYLER, C. R. KIRMAN, *et al.* (2001). Proposed occupational exposure limits for select ethylene glycol ethers using PBPK models and Monte Carlo simulations. *Toxicological Sciences* **62**(1), 124-139.
- VAN der MOLLEN G. W., S. A. L. M. KOIJMAN and W. SLOB (1996). A generic toxicokinetic model for persistent lipophilic compounds in humans : an application to TCDD. *Fundamental and Applied Toxicology* **31**, 83-94.
- WAKEFIELD J. and J. BENNETT (1996). The Bayesian modeling of covariates for population pharmacokinetic models. *Journal of the American Statistical Association* **91**, 917-927.
- WAKEFIELD J. and S. WALKER (1997). Bayesian nonparametric population models : formulation and comparison with likelihood approaches. *Journal of Pharmacokinetics and Biopharmaceutics* **25**(2), 235-53.
- WANG X., M. J. SANTOSTEFANO, M. J. DEVITO, *et al.* (2000). Extrapolation of a PBPK model for dioxins across dosage regimen, gender, strain, and species. *Toxicological Sciences* **56**, 49-60.
- WATANABE K. H. and F. Y. BOIS (1996). Interspecies extrapolation of physiological pharmacokinetic parameter distributions. *Risk Analysis : an Official Publication of the Society for Risk Analysis* **16**(6) : 741-754.