



HAL
open science

Benchmark dose : définitions, intérêt et usages en évaluation des risques sanitaires

Nathalie Bonvallot, Laurent Bodin, Cédric Duboudin, Denis Bard

► To cite this version:

Nathalie Bonvallot, Laurent Bodin, Cédric Duboudin, Denis Bard. Benchmark dose : définitions, intérêt et usages en évaluation des risques sanitaires. *Environnement, Risques & Santé*, 2009, 8 (6), pp.529-537. hal-00672491

HAL Id: hal-00672491

<https://hal.science/hal-00672491>

Submitted on 7 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Benchmark dose : définitions, intérêt et usages en évaluation des risques sanitaires

NATHALIE BONVALLOT¹
LAURENT BODIN²
CÉDRIC DUBOUDIN³
DENIS BARD¹

¹ École des hautes études
en santé publique
Avenue du professeur
Léon Bernard
CS 74312,
35043 Rennes cedex
<nathalie.bonvallot@ehesp.
fr>
<denis.bard@ehesp.fr>

² Agence française de sécurité
sanitaire de l'environnement
et du travail
253, avenue
du Général Leclerc
94701 Maisons-Alfort Cedex
<laurent.bodin@afsses.fr>

³ Union des caisses nationales
de sécurité sociale
18, avenue Léon Gaumont
75980 Paris cedex
<cduboudin@ucanss.fr>

Tirés à part :
N. Bonvallot

Article reçu le 11 mai 2009,
accepté le 2 novembre 2009

Résumé. Dans la démarche d'évaluation des risques sanitaires, les valeurs toxicologiques de référence (VTR) concernant les substances chimiques considérées comme ayant un seuil de dose sont traditionnellement construites à partir d'une dose critique, le plus souvent correspondant à la dose maximale sans effet néfaste observé (en anglais, le *No Observed Adverse Effect Level*, NOAEL) chez l'animal ou chez l'homme et l'application de facteurs d'incertitude. De nombreuses critiques ont été formulées sur l'utilisation du NOAEL, ce qui a conduit à la proposition d'une nouvelle dose critique : il s'agit de la *benchmark Dose* (BMD), de plus en plus utilisée depuis le début des années 2000. Cette note présente la méthode de construction de la BMD. Ses domaines d'application, ses avantages et ses limites sont également discutés.

Mots clés : BMD ; BMDL ; dose critique ; évaluation des risques sanitaires ; modélisation ; VTR.

Abstract

Benchmark dose: definitions, utility and use in health risk assessment

The process of assessing health risks related to chemical compounds believed to have a threshold dose traditionally relies on the derivation of a Toxicity Reference Value (TRV), based most often on the highest no-observed-adverse-effect level (NOAEL) found in experimental animal data, after the application of uncertainty factors. Criticism of the NOAEL approach over the years has led to an alternative method: the benchmark dose (BMD), which is used increasingly today. In this paper we present the methods used to construct a BMD and discuss its fields of application, its advantages and its limitations.

Key words: BMD; BMDL; critical dose; dose-response models; Human Toxicity Values; risk assessment.

Dans la démarche d'évaluation des risques sanitaires, pour chaque substance toxique identifiée, la comparaison des expositions de la population à des valeurs toxicologiques de référence (VTR) permet d'apprécier l'apparition d'une pathologie donnée, de manière qualitative (présence ou absence) ou quantitative (probabilité), en fonction des hypothèses mécanistiques sous-jacentes (mécanisme à seuil ou sans seuil de dose). Il s'agit de l'étape de caractérisation des risques [1]. Ces VTR sont élaborées pour chaque substance et pour différentes voies et périodes d'exposition (chronique ou aiguë, par voie orale ou respiratoire...), à partir des relations dose-réponses caractérisant l'apparition d'un effet (généralement l'effet critique, c'est-à-dire celui qui apparaît aux plus faibles expositions retrouvées dans la littérature). L'analyse de cette relation dose-réponse permet généralement de définir une dose critique qui sert de « point de départ » pour la construction de la VTR [2].

Les descripteurs de doses les plus connus et largement utilisés depuis plus de 40 ans sont les « NOAEL » et « LOAEL » (pour *No* ou *Lowest Observed Adverse Effect Level*), en français, DSEO

(dose maximale sans effet néfaste observé) et DMEO (dose minimale avec effet néfaste observé) [3]. Ils sont déterminés à partir de l'expérimentation animale, grâce à un test statistique qui permet de détecter une différence significative dans les effets ou la réponse induits par une dose testée par rapport à ceux qui sont observés sur un groupe non exposé (témoin) (*figure 1*) :

– le LOAEL correspond à la première dose testée, dans l'ordre croissant, qui produit un effet – considéré comme néfaste – d'une intensité ou d'une fréquence présentant une différence statistiquement significative avec le témoin. C'est donc la première dose du protocole pour laquelle on rejette l'hypothèse nulle (égalité de réponse entre la dose D_1 et le témoin). Le LOAEL ne correspond pas à un seuil biologique mais à un seuil statistique qui dépend du protocole expérimental ;

– le NOAEL est défini comme la dose testée immédiatement inférieure au LOAEL. Lorsque la première dose testée (D_1) conduit déjà à un effet significativement différent du témoin, il n'est pas possible de définir un NOAEL.

Par extension, en évaluation des risques, on parle également de NOAEL ou de LOAEL identifiés chez l'homme, même si

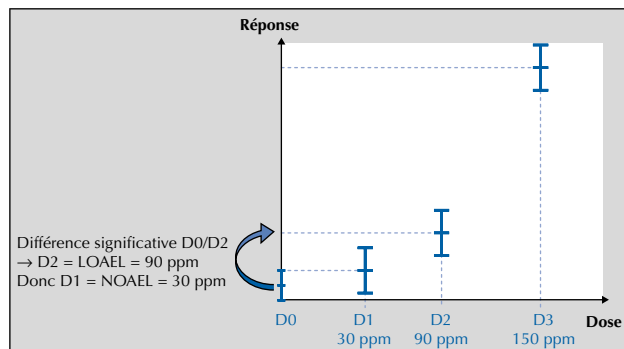


Figure 1. Exemple de résultats d'une étude expérimentale chez l'animal : mise en jeu d'un test statistique permettant de détecter une modification de réponse significative entre la dose testée et le groupe témoin.

LOAEL : Lowest Observed Adverse Effect Level ; NOAEL : No Observed Adverse Effect Level.

Figure 1. Example of results of an experimental animal study – application of a statistical test making it possible to detect a significant modification of response between the dose tested and the control group.

LOAEL : Lowest Observed Adverse Effect Level ; NOAEL : No Observed Adverse Effect Level.

l'exposition étudiée ne fait pas l'objet d'un protocole expérimental contrôlé. Dans ce cas, le LOAEL correspond à une exposition associée à un effet sanitaire dans la population étudiée, alors que le NOAEL correspond à une exposition n'ayant pas été associée à la présence d'un effet (c'est souvent l'exposition dans le groupe témoin qui est considérée comme un NOAEL).

Ces descripteurs de dose ont fait l'objet de nombreuses critiques [4, 5] :

- ils correspondent à une dose du protocole et sont donc très dépendants des choix initiaux de l'opérateur, ce qui peut conduire à une grande variabilité d'une étude à l'autre ;
- leur identification s'effectue sans exploiter l'ensemble de la relation dose-réponse ;
- leur usage n'encourage pas l'utilisation de protocoles puissants et est plutôt contraire au principe de prudence utilisé dans des situations de grande incertitude : moins le protocole est puissant et plus la dose sans effet apparent est élevée [6].

Ces limites, ainsi que la volonté, depuis quelques années, d'harmoniser les méthodes de construction des VTR, ont conduit à proposer de nouveaux descripteurs. L'un des plus étudiés et utilisés aujourd'hui est la *benchmark dose*, en français « dose repère », introduite par Crump en 1984 [4]. Actuellement, les agences nationales et internationales d'évaluation des risques, leurs comités d'experts, la Commission européenne et bien d'autres proposent, au travers de leurs lignes directrices ou de leurs avis, l'usage préférentiel de la BMD pour l'évaluation d'un risque [7-16]. Ceci est en accord avec la pratique : par exemple, l'exploitation de la base IRIS de l'US EPA montre que les premiers usages datent de la fin des années 1990 (figure 2) et que, depuis le début des années 2000, ce sont environ 40 à 100 % des valeurs de VTR à seuil (RfD [Reference Dose] et RfC [Reference Concentration]) produites chaque année qui sont élaborées à partir d'une *benchmark dose* (tableau 1). Enfin, en novembre 2008, la communauté internationale précisait à nouveau l'intérêt de l'utilisation de la *benchmark dose* à la place du NOAEL pour l'évaluation des risques sanitaires [17].

Dans ce contexte où le développement méthodologique est important et où l'intérêt de l'application de cette approche apparaît faire *consensus* au sein de la communauté scientifique internationale, il est opportun de proposer un document synthétique sur la *benchmark dose*. L'objectif de cet article est de présenter

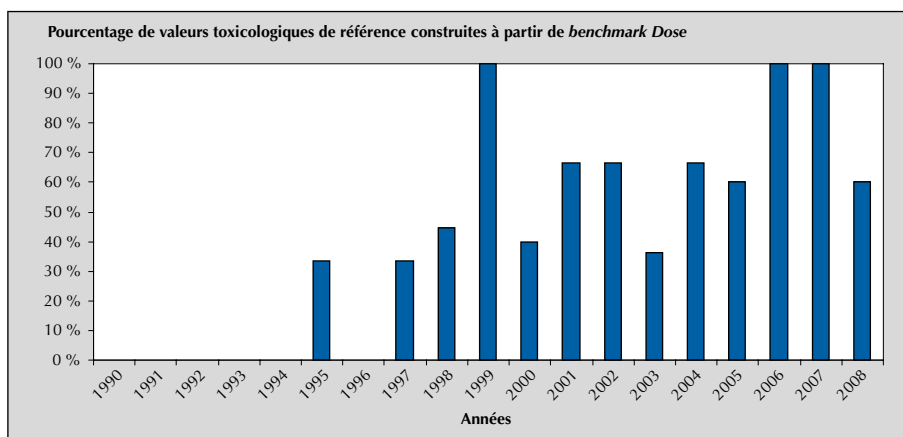


Figure 2. Pourcentage de valeurs toxicologiques de référence (VTR) construites annuellement par l'U.S.Environmental Protection Agency à partir d'une BMDL (limite inférieure de l'intervalle de confiance de la *benchmark dose*) de 1990 à 2008, par rapport au nombre total de VTR élaborées.

Extraction base IRIS, 9 décembre 2008.

Figure 2. Substances with a TRV constructed from a BMDL in the US EPA IRIS database.

Extracted from the US EPA IRID database, 9 December 2008.

Tableau 1. Substances possédant une valeur toxicologique de référence (VTR) construite à partir d'une BMDL (limite inférieure de l'intervalle de confiance de la *benchmark dose*), et présentes dans la base IRIS de l'*US Environmental Protection Agency* (extraction base 9 décembre 2008).

Table 1. Substances with a TRV constructed from a BMDL in the US EPA IRIS database (extracted 9 December 2008)

Numéro CAS	Nom de la substance	Type de VTR	Benchmark dose	Date
56-35-9	Oxyde de tributylétain	RfD	BMDL ₁₀ : 0,03 mg/kg/j	1997
67-66-3	Chloroforme	RfD	BMDL ₁₀ : 1 mg/kg/j	2001
71-43-2	Benzène	RfD	BMDL _{1SD} : 1,2 mg/kg/j	2003
71-43-2	Benzène	RfC	BMCL _{1SD} : 8,2 mg/m ³	2003
71-55-6	1,1,1-Trichloroéthane	RfD	BMDL ₁₀ : 2155 mg/kg/j	2007
71-55-6	1,1,1-Trichloroéthane	RfD	BMDL ₁₀ : 2155 mg/kg/j	2007
75-15-0	Disulfure de carbone	RfC	BMCL ₁₀ (HEC) : 19,7 mg/m ³	1995
75-35-4	1,1-Dichloroéthylène	RfD	BMDL ₁₀ : 4,6 mg/kg/j	2002
75-35-4	1,1-Dichloroéthylène	RfC	BMCL ₁₀ (HEC) : 6,9 mg/m ³	2002
75-44-5	Phosgène	RfD	BMDL ₁₀ (HEC) : 0,03 mg/m ³	2006
77-47-4	Hexachlorocyclopentadiène	RfD	BMDL ₁₀ : 6 mg/kg/j	2001
80-62-6	Métacrylate de méthyle	RfC	BMCL ₁₀ (HEC) : 7,2 mg/m ³	1998
91-57-6	2-Méthyl-naphthalène	RfD	BMDL ₀₅ : 4,7 mg/kg/j	2003
101-68-8	Diisocyanate de méthylène diphenyle monomère	RfC	BMCL ₁₀ (HEC) : 0,06 mg/m ³	1998
106-93-4	1,2-Dibromoéthane	RfC	BMCL ₁₀ (HEC) : 2,8 mg/m ³	2004
106-99-0	1,3-Butadiène	RfC	BMCL ₁₀ (HEC) : 1,98 mg/m ³	2002
108-88-3	Toluène	RfD	BMDL _{1SD} : 238 mg/kg/j	2005
108-95-2	Phénol	RfD	BMDL _{1SD} : 93 mg/kg/j	2002
110-54-3	n-Héxane	RfC	BMCL _{1SD} (HEC) : 215 mg/m ³	2005
110-82-7	Cyclohexane	RfC	BMCL _{1SD} (HEC) : 1822 mg/m ³	2003
111-76-2	2-butoxyéthanol	RfD	BMDL ₀₅ (HED) : 5,1 mg/kg/j	1999
111-76-2	2-butoxyéthanol	RfC	BMCL ₀₅ (HEC) : 380 mg/m ³	1999
123-38-6	Propionaldéhyde	RfC	BMCL ₁₀ (HEC) : 8 mg/m ³	2008
542-75-6	1,3-Dichloropropène	RfD	BMDL ₁₀ : 3,4 mg/kg/j	2000
542-75-6	1,3-Dichloropropène	RfC	BMCL ₁₀ (HEC) : 0,72 mg/m ³	2000
811-97-2	1,1,1,2-Tétrafluoroéthane	RfC	BMCL ₁₀ (HEC) : 8200 mg/m ³	1995
1309-64-4	Trixyde d'antimoine	RfC	BMCL ₁₀ (HEC) : 0,074 mg/m ³	1995
5436-43-1	2,2',4,4'-Tétra-bromodiphényléther	RfD	BMDL _{1SD} : 0,35 mg/kg/j	2008
7440-39-3	Barium et composés	RfD	BMDL ₀₅ : 63 mg/kg/j	2005
7440-41-7	Béryllium et composés	RfD	BMDL ₁₀ : 0,46 mg/kg/j	1998
7440-42-8	Bore et composés	RfD	BMDL ₀₅ : 10,3 mg/kg/j	2004
7664-38-2	Acide phosphorique	RfC	BMCL ₁₀ (HEC) : 3,4 mg/m ³	1995
9016-87-9	Diisocyanate de méthylène diphenyle polymère	RfC	BMCL ₁₀ (HEC) : 0,06 mg/m ³	1998
18540-29-9	Chrome VI	RfC	BMCL ₁₀ (ADJ) : 0,034 mg/m ³	1998
22967-92-6	Méthylmercure	RfD	BMDL ₀₅ : 0,001472 mg/kg/j	2001
22967-92-6	Méthylmercure	RfD	BMDL ₀₅ : 0,000857 mg/kg/j	2001
60348-60-9	2,2',4,4',5-Pentabromodiphényléther	RfD	BMDL _{1SD} : 0,29 mg/kg/j	2008

cette approche, les usages de la *benchmark dose* en évaluation des risques, ses intérêts et ses limites.

Définition de la *benchmark dose*

La *benchmark dose* est une dose produisant un effet non nul, correspondant à une modification du niveau de réponse par rapport à un groupe témoin. Le plus souvent, ce n'est pas directement la *benchmark dose* qui est utilisée, mais la limite inférieure de son intervalle de confiance à 95 % (BMDL). Par commodité, on utilisera dans la suite de cet article « BMD » comme un acronyme générique pour évoquer la *benchmark dose* ou la

limite inférieure de son intervalle de confiance. Le principe de cette approche repose sur une modélisation des données expérimentales prenant en compte l'ensemble de la relation dose-réponse (figure 3). L'analyse de la réponse dans les groupes exposés nécessite qu'on définisse plusieurs cas de figure en fonction du type d'effets choisis :

- si les observations concernent le nombre d'animaux atteints (par une altération organique, une pathologie, etc.), la réponse obtenue est dite dichotomique ou quantale : c'est la proportion d'individus touchés ;
- si les observations concernent un paramètre physiologique ou biologique de l'organisme (par exemple, le poids d'un

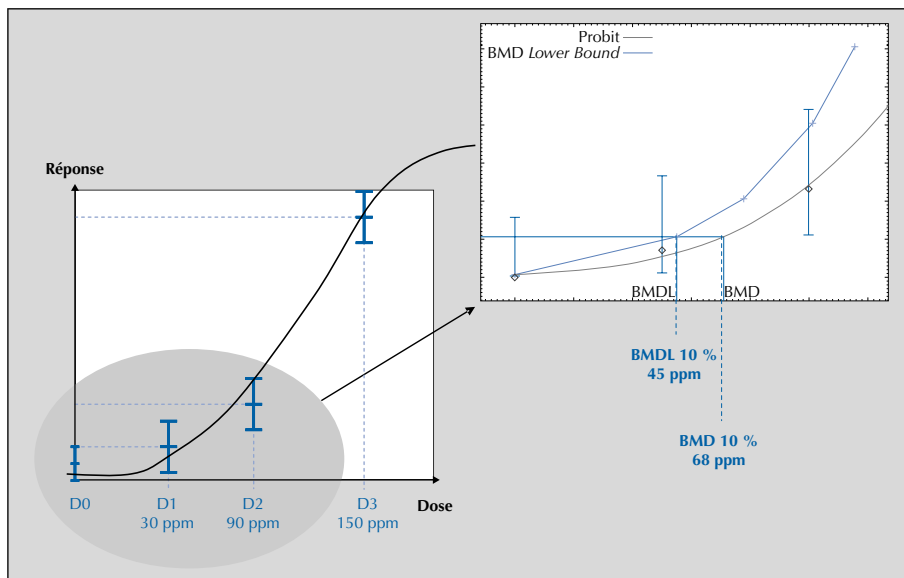


Figure 3. Exemple de construction d'une benchmark dose (BMD et BMDL) à partir de données expérimentales.

Figure 3. Example of the construction of benchmark doses (BMD and BMDL) based on experimental data.

organe, le nombre de globules rouges ou la concentration sanguine d'une enzyme hépatique), la réponse obtenue est dite de nature continue.

En fonction de la nature de la réponse, les modèles ou équations mathématiques utilisées pour la construction de la BMD ne seront pas les mêmes.

Réponse dichotomique

Dans le cas de données dichotomiques, la $BMD_{x\%}$ est définie comme la dose qui engendre une augmentation de $x\%$ de l'incidence de la réponse observée dans le groupe exposé par rapport au témoin. Cette augmentation de $x\%$, appelée *benchmark response* (BMR) peut être définie de deux manières différentes, comme [13] :

- un écart absolu par rapport à l'incidence observée dans le groupe témoin : $BMR = I(BMD) - I(0)$ où $I(BMD)$ est l'incidence de la dose correspondant à la BMD et $I(0)$ l'incidence au sein du groupe témoin (on parle d'*additive risk* en anglais) ;
- un pourcentage de l'amplitude maximale de l'incidence observée dans le groupe témoin :

$$BMR = \frac{I(BMD) - I(0)}{1 - I(0)}$$

C'est un excès de risque standardisé (d'*extra risk*). Pour des raisons de comparabilité des données, cette approche est plus souvent utilisée.

Réponse continue

Dans le cas de données continues, la réponse BMR qui détermine la BMD peut également être définie de plusieurs manières [13] :

- 1) comme un écart absolu par rapport à la réponse observée sur le groupe témoin : $BMR = R(BMD) - R(0)$ où $R(BMD)$ est la réponse moyenne à la dose BMD et $R(0)$ est la réponse moyenne chez les témoins (« *Absolute Change in Mean* ») ;
- 2) comme un écart relatif par rapport à la réponse observée chez le groupe témoin rapporté à cette réponse (« *Relative Change in Mean* ») :

$$BMR = \frac{R(BMD) - R(0)}{R(0)}$$

- 3) comme un écart par rapport à la réponse observée chez le groupe témoin standardisé par l'écart type de cette réponse :

$$BMR = \frac{R(BMD) - R(0)}{SD(0)}$$

où $SD(0)$ est l'écart type du groupe témoin (« *Change in Mean Relative to Standard Deviation of Control* ») ;

- 4) comme un écart par rapport à la réponse observée chez le groupe témoin standardisé par l'écart maximal observé sur les réponses :

$$BMR = \frac{R(BMD) - R(0)}{R() - R(0)}$$

où $R()$ est la plus grande réponse observée, lorsque la dose tend vers l'infini (« *Change beyond Background Standardized by Total Range of Response* »).

Suivant le choix effectué, la BMR est donc de nature et d'unité différente. Au final, il s'agit de choisir une valeur numérique pour cette BMR, à partir de laquelle on considère la réponse observée comme anormale (pour illustration, voir l'exemple du toluène) et qui détermine la BMD. Le choix de la relation entre BMD et BMR doit être clairement expliqué.

Intérêt de l'utilisation de la benchmark dose

L'approche BMD est aujourd'hui développée, car elle comporte un certain nombre d'avantages par rapport à l'approche traditionnelle :

- l'utilisation de la limite inférieure de l'intervalle de confiance à 95 % de la BMD permet de prendre en compte l'incertitude liée à la réalisation de l'expérimentation et la variabilité interindividuelle (qui est généralement faible chez l'animal). On choisit la borne inférieure de cet intervalle de confiance pour respecter le principe de prudence, d'autant plus important en cas de grande incertitude ;

- la prise en compte de l'ensemble de la relation dose-réponse, en considérant toutes les données de l'expérimentation animale, entraîne une variabilité sur la BMD beaucoup plus faible que celle que l'on peut trouver sur un NOAEL ou un LOAEL. Cela a été montré à partir de l'analyse empirique de données concernant à la fois les NOAEL et les BMD [18, 19] ;

- la modélisation de la relation dose-réponse permet d'associer un niveau de risque (excès de risque) à chaque dose (ou exposition) ce qui a un grand intérêt en termes décisionnels ;

- contrairement au NOAEL, l'utilisation de la BMDL pénalise le protocole expérimental et l'expérimentation de moins bonne qualité, car elle conduit à dériver une valeur plus faible, ce qui est en accord avec le principe de prudence évoqué plus haut [8, 20].

Ainsi, l'approche BMD permet de rendre explicites des choix ou hypothèses (niveau de réponse ou d'effet jugé comme néfaste par exemple) bien souvent implicites dans l'approche NOAEL/LOAEL.

Enfin, la disponibilité de descripteurs de dose comparables entre eux permet de répondre à un objectif de hiérarchisation très utile en termes de santé publique.

Méthodes et utilisations

Des outils tels que BMD (pour Benchmark dose) Software de l'US EPA (US Environmental Protection Agency) et PROAST du RIVM (*National Institute for the Public Health and the Environment*) ont été développés ces dernières années spécifiquement pour l'estimation des BMD. Plusieurs étapes sont nécessaires, en particulier l'ajustement et le choix d'un modèle théorique de relation dose-réponse aux données expérimentales, et le choix du niveau de réponse à utiliser.

Ajustement de modèles sur les données expérimentales et sélection

En pratique, les modèles mathématiques ou statistiques sont souvent plus utilisés que les modèles biologiques, car ils présentent un nombre de paramètres plus faibles. Néanmoins, l'utilisation de modèles statistiques doit être en accord avec les phénomènes biologiques connus liés aux substances étudiées [21]. Il est habituel d'ajuster plusieurs modèles aux données expérimentales, ce qui nécessite un choix en seconde étape. Les modè-

les utilisés sont des modèles fondés sur des distributions de probabilité ou des modèles multi-étapes, fondés sur l'hypothèse que la réponse observée chez un individu exposé est le résultat de la mise en jeu (aléatoire ou en ordre spécifié) d'un ou plusieurs événements biologiques. Plusieurs revues de la littérature ont largement étudié l'application de ces modèles [22, 23].

L'ajustement aux données expérimentales, qui consiste à déterminer les valeurs des paramètres du modèle, peut être fait à partir de la méthode du maximum de vraisemblance [23]. Les valeurs retenues sont celles qui maximisent la probabilité d'occurrence des valeurs expérimentales observées. Ensuite, l'utilisation d'un test d'ajustement comme le test du Chi² de Pearson permet de vérifier si le modèle rend correctement compte des données.

Enfin, le choix d'un modèle parmi plusieurs testés doit être réalisé après l'observation graphique (analyse visuelle) du modèle et l'analyse des résidus à partir de la méthode du maximum de vraisemblance. Le critère d'Akaike ou AIC (*Akaike Information Criterion*) est classiquement utilisé pour identifier le « meilleur » modèle ; son principe consiste à identifier un optimum entre la capacité du modèle à « coller » aux données et le nombre de paramètres du modèle. Cette procédure est par exemple implémentée dans le logiciel de l'US EPA. Le critère AIC mesure l'adéquation des modèles aux données expérimentales, à partir de la robustesse de l'ajustement ou du maximum de vraisemblance et du nombre de paramètres [24]. Le modèle qui a alors la valeur de l'AIC la plus faible est retenu.

Choix de la réponse associée à la BMD (BMR)

Le choix de la valeur numérique de la BMR diffère en fonction de la nature et de la qualité des données disponibles. Dans le cas d'une réponse dichotomique, le choix du pourcentage associé à la BMD (proportion d'individus) se fait en fonction de la sensibilité du test expérimental, de la sévérité de l'effet étudié et de l'incidence de base retrouvée chez les témoins. Dans la pratique, des valeurs par défaut de 1,5 et 10 % sont plus souvent employées et recommandées par l'US EPA [25]. Ces valeurs sont fondées sur la comparaison de NOAEL et de BMDL pour plus de 200 études expérimentales de toxicité du développement [26, 27], qui a montré que les BMDL 1 % et 5 % (en excès de risque standardisé) étaient généralement bien inférieures aux NOAEL, que la limite de détection dans les études traditionnelles de toxicité prénatale était de 5 % et que les BMDL₁₀ étaient le plus souvent très proches des NOAEL. Par ailleurs, d'autres travaux ont montré que la valeur de 10 % représentait la limite de sensibilité des essais toxicologiques standardisés de cancérogenèse ou de toxicité à doses répétées [28, 29]. Ces observations sont à la base des recommandations pour le choix de la BMR de 5 ou 10 % en fonction des cas. Néanmoins, ces recommandations ne sont pas fondées sur des aspects strictement scientifiques. Elles reposent sur les niveaux de risque associés aux NOAEL traditionnellement utilisés depuis les années 1950 [23].

Dans le cas de données continues, la principale difficulté concerne le choix de la modification maximale que l'on tolère comme étant physiologique (ou non néfaste) dans le paramètre étudié. Cela revient à se demander quelle variation biologique peut être considérée comme acceptable d'un point de vue phy-

Tableau 2. Résultats de l'étude des effets du toluène sur le poids de naissance chez le rat Sprague-Dawley exposé de 0 à 2 000 ppm [35].

Table 2. Results of the study of the effects of toluene on birth weight in Sprague-Dawley rats exposed to 0 to 2000 ppm [35].

Exposition (ppm)	Génération F1					
	Mâles			Femelles		
	Nombre d'animaux examinés	Moyenne des poids	Écart type des poids	Nombre d'animaux examinés	Moyenne des poids	Écart type des poids
0	188	6,50	0,58	167	6,00	0,58
100	103	6,40	0,55	100	5,90	0,59
500	109	6,40	0,65	109	6,20	0,81
2 000	88	5,70*	0,63	83	5,30*	0,69

siologique. Cette question oblige l'évaluateur à définir un niveau de réponse considéré comme « néfaste », à partir d'arguments descriptifs biologiques, toxicologiques et statistiques, et par l'analyse des données sur les témoins historiques. Différents auteurs se sont penchés sur cette problématique [7, 21, 30-33]. Dans l'idéal, ce seuil devrait être fondé sur la signification biologique de la réponse étudiée. Une réponse plus faible peut, par exemple, être retenue dans le cas d'effets particulièrement sévères, ou encore une valeur plus élevée peut être proposée lorsque l'étude porte sur un paramètre biologique très variable d'un individu à l'autre (par exemple, le niveau de transaminases sériques) [9], mais cela est difficilement atteignable dans la démarche d'évaluation des risques. Crump (1995) et Dekkers (2001) considèrent comme seuil la valeur de la moyenne du paramètre estimé chez les témoins, plus ou moins une fois l'écart type de cette valeur chez les témoins de l'expérimentation ou chez les témoins historiques [30, 33]. Cette approche a été proposée par l'US EPA comme approche par défaut, lorsqu'aucun élément n'est disponible [7]. Gaylor (2004) proposait d'utiliser les valeurs des quantiles 1 % et 99 % [31]. Plus récemment, Dekkers et ses collaborateurs ont estimé la variabilité interindividuelle et temporelle chez les animaux de laboratoire sains et non exposés pour plusieurs paramètres biologiques habituellement mesurés dans les études toxicologiques et pouvant constituer une réponse continue, en particulier des paramètres hématologiques (hémoglobine, nombre de globules rouges, volume corpusculaire moyen...) ou cliniques (transaminases, albumine, triglycérides...), en partant du principe que ces variations pouvaient être considérées comme « normales ». Ils ont mis en évidence que ces variations pouvaient être très différentes d'un paramètre à l'autre (de 0,6 à 34 %), et souvent éloignées des valeurs de 5 ou 10 % conventionnellement prises par défaut [34]. Finalement, le choix de la réponse repose avant tout sur un dialogue entre le toxicologue et le statisticien afin de retenir les options qui se rapprochent le plus de la réalité biologique connue au moment de l'évaluation.

Exemple du toluène

Afin d'illustrer les difficultés rencontrées dans le choix de la BMR pour des données continues, l'exemple de la construction de plusieurs BMDL pour le toluène, à partir du logiciel de l'US EPA (BMDS version 1.4.1c) est développé ci-dessous.

L'étude de Roberts *et al.* 2003 décrit les effets liés à une exposition au toluène sur plusieurs générations [35]. Le *tableau 2* résume les résultats observés chez les rats Sprague Dawley de la génération F1 exposés à 0, 100, 500 et 2 000 ppm par inhalation. Dans cette étude, la diminution du poids à la naissance a été retenue comme effet critique. La diminution statistiquement significative du poids des nouveau-nés mâles et femelles des générations F1 a été observée à la dose de 2 000 ppm, ce qui correspond donc au LOAEL. Le NOAEL est par définition situé à 500 ppm.

Pour construire la BMDL, nous avons utilisé plusieurs approches. Celle de l'US EPA, préconisée par défaut, correspond au calcul de la BMD par la formule de l'écart relatif à l'écart type avec une valeur de BMR = -1, soit la formule suivante [7] : $R(BMD) = R(0) - SD(0)$

Dans ces conditions, et selon les recommandations de l'US EPA concernant le choix du modèle final¹, les valeurs de BMD et de BMDL obtenues sont respectivement de 1 715 ppm et de 1 420 ppm² (autrement dit, et à titre de comparaison, cela revient à considérer comme réponse « anormale » une diminution de 9 % du poids par rapport au groupe témoin).

L'approche proposée par Dekkers en 2006 correspond :

- au calcul de la BMR selon la formule suivante [34] : BMR = limite inférieure de l'IC 99 % du groupe témoin ;
- à une BMD par la formule : $R(BMD) = -BMR$

À partir des données du *tableau 2*, l'intervalle de confiance (IC 95 %) de la moyenne des poids corporels du groupe témoin de la génération F1 a été calculé. Ainsi, pour un nombre d'animaux égal à 188, une moyenne de 6,50 grammes et un écart type de 0,58 gramme, on obtient un IC 95 % compris entre 6,39 et 6,61 grammes, ce qui revient à considérer comme BMR tout poids de nouveau-né inférieur à 6,39 grammes pour la génération F1 – autrement dit, et à titre de comparaison, cela revient à considérer comme réponse « anormale » une diminution de 1,6 % du poids par rapport au groupe témoin. Les valeurs de

¹ Les recommandations et la démarche permettant de sélectionner le modèle final ne seront pas développées dans cette note. Pour plus d'informations, le lecteur peut se rendre sur le site de l'US EPA : <http://www.epa.gov/ncea/bmds/>.

² C'est le modèle de puissance qui a été sélectionné pour la modélisation de la courbe dose-réponse.

BMD et de BMDL ainsi obtenues sont respectivement de 502 ppm et de 231 ppm (Bonvallot *et al.*, en préparation).

Cet exemple montre que c'est le niveau de réponse choisi (dans un cas 9 %, dans l'autre 1,6 %) qui influence les résultats de la BMD et de la BMDL, plus que le modèle mathématique lui-même – cela n'a pas été détaillé ici, mais le choix d'un autre modèle, tel que le modèle de Hill aurait conduit à des résultats du même ordre de grandeur. À ce stade il n'y a pas d'arguments pour préférer un niveau-réponse à un autre, mais cette étude montre qu'il n'y a pas plus d'arguments à retenir un NOAEL. En effet, dans ce cas, le NOAEL correspond à une dose possédant un effet sur la diminution du poids corporel d'au moins 1,5 %.

Limites

Les limites associées à la construction de BMD sont aujourd'hui essentiellement liées à la plus grande complexité méthodologique de cette approche par rapport à l'approche traditionnelle utilisant le NOAEL. Cet aspect est responsable d'une certaine réticence à utiliser l'approche BMD, en particulier dans le contexte de l'évaluation des risques comme outil d'aide à la décision (36). En effet, l'utilisation de la BMD nécessite de voir posées explicitement des questions sur le choix d'un modèle d'ajustement de relation dose-réponse et le choix d'un niveau de réponse jugé comme acceptable d'un point de vue biologique et statistique, en particulier lors de l'utilisation de réponse continue. De plus, les compétences sont encore aujourd'hui peu développées et la pratique est relativement récente, ce qui ne permet pas un retour d'expérience très important. Les aspects méthodologiques et pratiques sont d'ailleurs toujours beaucoup discutés dans la littérature actuelle [23, 34, 36-43].

Historiquement, l'évaluation des risques est une approche qui a été proposée pour aider le gestionnaire dans un contexte d'incertitude scientifique. Il fallait donc être capable de proposer des VTR dans des contextes de connaissances très limitées. Cela a été rendu possible avec l'approche traditionnelle, puisque, lors de l'utilisation de données animales peu fournies, il a été proposé de retenir le LOAEL lorsque la première dose testée conduit déjà à une réponse significativement différente du témoin. De même, lorsqu'aucun effet n'est observé dans une étude, il a été proposé de retenir la dernière dose testée comme étant un NOAEL. Pour l'approche BMD, la modélisation des données nécessite *a priori* de disposer de plusieurs groupes de doses encadrant l'absence et la présence de l'effet : l'ajustement n'en sera que plus précis. Si un LOAEL est observé à la première dose testée (pas de NOAEL), la construction d'une BMD engendre une extrapolation en dehors du domaine des observations, ce qui peut entraîner des divergences entre plusieurs BMD construites à partir de plusieurs modèles, sans qu'un choix éclairé puisse être fait. Si aucun effet n'a été observé (pas de LOAEL), la tendance de la relation dose-réponse ne peut pas être connue et une BMDL ne peut pas être proposée [23]. Murrell *et al.* ont remarqué que, dans le cas d'un protocole expérimental peu puissant, la BMDL peut être inférieure de plus d'un ordre de grandeur à la BMD, ce qui interroge sur la vraisemblance de son utilisation [44]. Nous l'avons également mis en évidence dans le cas du di-n-butylphthalate [45].

Enfin, il faut également noter que la mise en application de l'approche BMD engendre une communication plus difficile en évaluation des risques, car elle nécessite de passer d'une approche « sans effet statistiquement significatif » par la définition d'un NOAEL à une approche « avec effet » dont on choisit la réponse *a priori*. Les travaux montrant que les NOAEL ne sont pas réellement des doses sans effet [26, 27] ne suffisent pas à rendre la communication plus facile.

Conclusion et perspectives

Cet article a permis de réaliser un état des lieux des connaissances relatives aux méthodes de construction des *benchmark doses* et à leurs usages dans la démarche d'évaluation des risques. Cette synthèse montre que l'usage des BMD nécessite des choix méthodologiques concernant les modèles de relations dose-réponse, les ajustements et le niveau de réponse associé à la BMD. Ces choix ne font pas encore *consensus* actuellement dans la communauté des évaluateurs de risques, notamment parce que le retour d'expérience pratique est encore insuffisant. Néanmoins, la littérature actuelle est riche et de nombreux auteurs ont proposé des pistes de réflexion sur la manière de réaliser ces choix. Encore une fois, la transparence est de mise dans cette démarche.

La BMD dispose probablement d'un avenir florissant et devrait être de plus en plus utilisée, compte tenu de ses avantages dans bien des domaines (diminution des incertitudes, approche probabiliste de l'évaluation des risques sanitaires, approches comparatives en santé publique, principe de précaution...).

L'une des principales perspectives concerne la validité des protocoles d'études expérimentaux actuels pour la construction de BMD. Ces protocoles comprennent typiquement un groupe témoin et trois groupes de doses qui doivent être choisis en fonction de données préalables concernant les effets des substances. Un nombre fixé d'animaux par groupe de doses est également défini (par exemple, dans une étude de toxicité chronique chez le rongeur, il doit y avoir au moins 20 animaux par groupe de dose et par sexe) [46]. La plupart des analyses bibliographiques effectuées ces dernières années montrent que ces protocoles d'études sont *a priori* suffisants pour la définition de la BMD mais qu'il semble préférable de privilégier un plus grand nombre de groupes de doses sans augmenter le nombre d'animaux traités [23, 41, 47-49].

Enfin, l'utilisation de la BMD laisse aujourd'hui de belles perspectives concernant l'application de nouveaux outils de toxicologie en évaluation des risques. Les exemples décrits à partir de l'exploitation de données de génomique montrent qu'il est possible de modéliser les modifications moléculaires et cellulaires associées à l'exposition à des substances, permettant la construction de VTR fondée sur l'expression génique [50, 51]. ■

Remerciements et autres mentions

Financement : aucun ; conflit d'intérêt aucun.

Références

1. Williams P, Paustenbach D. Risk characterization. In : Paustenbach DJ, ed. *Human and Ecological Risk Assessment: theory and practice*. Chichester : Wiley-Interscience, 2002.
2. Bonvallot N, Dor F. Analyse des méthodes d'élaboration des VTR: une aide à la sélection? *Environnement, Risques et Santé* 2002 ; 1 : 178-83.
3. Bard D. Principes de l'évaluation des risques pour la santé publique liés aux expositions environnementales. *Rev Epidemiol Sante Publique* 1995 ; 43 : 423-31.
4. Crump KS. A new method for determining allowable daily intakes. *Fundam Appl Toxicol* 1984 ; 4 : 854-71.
5. Gaylor D, Ryan L, Krewski D, Zhu Y. Procedures for Calculating Benchmark Doses for Health Risk Assessment. *Regul Toxicol Pharmacol* 1998 ; 28 : 150-64.
6. Bois FY. NOAEL - LOAEL. Mot à mot. *Environnement, Risques et Santé* 2008 ; 7 : 441.
7. U.S. Environmental Protection Agency. *Benchmark dose technical guidance document. Externa Review Draft. Report No.: EPA/630/R-00/001*. Washington DC : U.S. Environmental Protection Agency, 2000.
8. Health Council of the Netherlands. *Benchmark dose method: derivation of health-based recommended exposure limits in new perspective*. Report No.: 2003/06E. La Hague : Health Council of the Netherlands, 2003.
9. California Environmental Protection Agency Department of Pesticide Regulation. *Guidance for Benchmark Dose (BMD) Approach – Continuous Data. Report No.: DPR MT-2*. Sacramento : California Environmental Protection Agency Department of Pesticide Regulation, 2004.
10. California Environmental Protection Agency Department of Pesticide Regulation. *Guidance for Benchmark Dose (BMD) Approach - Quantal Data. Report No.: DPR MT-1*. Sacramento : California Environmental Protection Agency Department of Pesticide Regulation, 2004.
11. European Food Safety Authority. *Mandate proposed to EFSA by the Scientific Committee on the use of the benchmark dose approach in risk assessment. Report No.: EFSA-Q-2005-232*. Parme : European Food Safety Authority, 2005.
12. European Food Safety Authority. *Opinion of the Scientific Committee on a request from EFSA related to a harmonised approach for risk assessment of substances which are both genotoxic and carcinogenic. Request No.: EFSA-Q-2004-020*. Parme : European Food Safety Authority, 2005.
13. Agence française de sécurité sanitaire de l'environnement et du travail (Afsset). Valeurs toxicologiques de référence (VTR) pour les substances reprotoxiques. Méthode de construction de VTR fondées sur des effets toxiques pour la reproduction et le développement. Maisons-Alfort : Agence française de sécurité sanitaire de l'environnement et du travail, 2007.
14. Fitzgerald DJ, Robinson NI. Development of a tolerable daily intake for N-nitrosodimethylamine using a modified benchmark dose methodology. *J Toxicol Environ Health* 2007 ; 70 : 1670-8.
15. Scientific Committee on Health and Environmental Risks (SCHER), Scientific Committee on Consumer Products (SCCP), Scientific Committee on Emerging and Newly Identified Health Risks (SCENIHR). *Opinion of the Scientific Committees on Risk assessment methodologies and approaches for mutagenic and carcinogenic substances*. Brussels : Commission européenne, 2009 http://ec.europa.eu/health/ph_risk/committees/04_scher/docs/scher_o_113.pdf.
16. European Chemicals Agency. Characterisation of dose [concentration]-response for human health. In: *Guidance for the implementation of REACH. Guidance on information requirements and chemical safety assessment*. Helsinki: European Chemicals Agency, 2008.
17. Bolt HM. Assessing the risk of non-threshold carcinogens. First International Conference on Risk Assessment "A global Risk Assessment Dialogue". Parallel session 2. 13-14 November 2008, Bruxelles.
18. Bokkers BG, Slob W. A comparison of ratio distributions based on the NOAEL and the benchmark approach for subchronic-to-chronic extrapolation. *Toxicol Sci* 2005 ; 85 : 1033-40.
19. Bokkers BG, Slob W. Deriving a data-based interspecies assessment factor using the NOAEL and the benchmark dose approach. *Crit Rev Toxicol* 2007 ; 37 : 355-73.
20. Leisenring W, Ryan L. Statistical properties of the NOAEL. *Regul Toxicol Pharmacol* 1992 ; 15 : 161-71.
21. Slob W. Dose-response modelling of continuous endpoints. *Toxicol Sci* 2002 ; 66 : 298-312.
22. Parham F, Portier C. Benchmark Dose Approach. In: Edler L, Kistos C, eds. *Recent Advances in Quantitative Methods in Cancer and Human Health Risk Assessment*. [Chichester : John Wiley & Sons Ltd, 2005.
23. Sand S, Victorin K, Filipsson AF. The current state of knowledge on the use of the benchmark dose concept in risk assessment. *J Appl Toxicol* 2008 ; 28 : 405-21.
24. Akaike H. *Information Theory and an Extension of the Maximum Likelihood Principle*. Proceedings of the 2nd International Symposium on Information Theory. Budapest, 1973.
25. U.S. Environmental Protection Agency. *A Review of the Reference Dose and Reference Concentration processes. Report No.: EP/630/P-02/002F*. Washington DC : U.S. Environmental Protection Agency, 2002.
26. Faustman EM, Allen BC, Kavlock RJ, Kimmel CA. Dose-Response Assessment for Developmental Toxicity: I. Characterization of Database and Determination of No Observed Adverse Effect Levels. *Fundam Appl Toxicol* 1994 ; 23 : 478-86.
27. Allen BC, Kavlock RJ, Kimmel CA, Faustman EM. Dose-response assessment for developmental toxicity. II. Comparison of generic benchmark dose estimates with no observed adverse effect levels. *Fundam Appl Toxicol* 1994 ; 23 : 487-95.
28. Haag-Grönlund M, Fransson-Steen R, Victorin K. Application of the benchmark method to trichloroethene toxicity data. *Toxicol Lett* 1994 ; 74 : 33.
29. Filipsson AF, Victorin K. Comparison of available benchmark dose softwares and models using trichloroethylene as a model substance. *Regul Toxicol Pharmacol* 2003 ; 37 : 343-55.
30. Crump KS. Calculation of benchmark doses from continuous data. *Risk Anal* 1995 ; 15 : 79-89.
31. Gaylor DW, Slikker WJr. Role of the standard deviation in the estimation of benchmark doses with continuous data. *Risk Anal* 2004 ; 24 : 1683-7.
32. Slob W, Pieters MN. probabilistic approach for deriving acceptable human intake limits and human health risks from toxicological studies: general framework. *Risk Anal* 1998 ; 18 : 787-98.
33. Dekkers S, de Heer C, Rennen MAJ. Critical effect sizes in toxicological risk assessment: a comprehensive and critical evaluation. *Environ Toxicol Pharmacol* 2001 ; 10 : 33-52.

34. Dekkers S, Telman J, Rennen MA, Appel MJ, de HC. Within-animal variation as an indication of the minimal magnitude of the critical effect size for continuous toxicological parameters applicable in the benchmark dose approach. *Risk Anal* 2006 ; 26 : 867-80.
35. Roberts LG, Bevans AC, Schreiner CA. Developmental and reproductive toxicity evaluation of toluene vapor in the rat. I. Reproductive toxicity. *Reprod Toxicol* 2003 ; 17 : 649-58.
36. Travis KZ, Pate I, Welsh ZK. The role of the benchmark dose in a regulatory context. *Regul Toxicol Pharmacol* 2005 ; 43 : 280-91.
37. Budtz-Jorgensen E. Estimation of the benchmark dose by structural equation models. *Biostatistics* 2007 ; 8 : 675-88.
38. Wheeler MW, Bailer AJ. Properties of model-averaged BMDLs: a study of model averaging in dichotomous response risk estimation. *Risk Anal* 2007 ; 27 : 659-70.
39. Zhu Y, Wang T, Jelsovsky JZ. Bootstrap estimation of benchmark doses and confidence limits with clustered quantal data. *Risk Anal* 2007 ; 27 : 447-65.
40. Piegorsch WW, West RW. Benchmark analysis: shopping with proper confidence. *Risk Anal* 2005 ; 25 : 913-20.
41. Kuljus K, von RD, Sand S, Victorin K. Comparing experimental designs for benchmark dose calculations for continuous endpoints. *Risk Anal* 2006 ; 26 : 1031-43.
42. Sand S, von RD, Victorin K, Filipsson AF. Identification of a critical dose level for risk assessment: developments in benchmark dose analysis of continuous endpoints. *Toxicol Sci* 2006 ; 90 : 241-51.
43. Nitcheva DK, Piegorsch WW, West RW, Kodell RL. Multiplicity-adjusted inferences in risk assessment: benchmark analysis with quantal response data. *Biometrics* 2005 ; 61 : 277-86.
44. Murrel JA, Protier C, Morris R. Characterizing dose-response I: critical assessment of the benchmark dose concept. *Risk Anal* 1998 ; 18 : 13-26.
45. Bonvallot N, Doornaert B, Dor F, et al. *Use of benchmark doses for reprotoxic chemicals*. Proceedings of the first International Conference on Fetal Programming and Developmental Toxicity (PPTOX), Torshavn, May 2007.
46. Organization for Economic Co-operation and Development. *Guideline for testing of chemical: chronic toxicity studies*. OECD Guideline n°452. Paris : OECD, 1981.
47. Kavlock RJ, Schmid JE, Setzer Jr RW. A simulation study of the influence of study design on the estimation of benchmark doses for developmental toxicity. *Risk Anal* 1996 ; 16 : 399-410.
48. Krewski D, Smythe R, Fung KY. Optimal designs for estimating the effective dose in developmental toxicity experiments. *Risk Anal* 2002 ; 22 : 1195-205.
49. Slob W, Moerbeek M, Rauniomaa E, Piersma AH. A statistical evaluation of toxicity study designs for the estimation of the benchmark dose in continuous endpoints. *Toxicol Sci* 2005 ; 84 : 167-85.
50. Thomas RS, Allen BC, Nong A, et al. A method to integrate benchmark dose estimates with genomic data to assess the functional effects of chemical exposure. *Toxicol Sci* 2007 ; 98 : 240-8.
51. Yang L, Allen BC, Thomas RS. BMDExpress: a software tool for the benchmark dose analyses of genomic data. *BMC Genomics* 2007 ; 8 : 387.