Provisional translation

Adapted October, 2019

Third edition (September, 2023)

The Food Safety Commission

This English version of the Expert Committee Decision is intended to be reference material to provide convenience for users. In the event of inconsistency between the Japanese original and this English translation, the former shall prevail. The Food Safety Commission of Japan (FSCJ) shall not be responsible for any consequence resulting from use of this English version.

Guidance on the Use of the BMD Approach in Risk Assessment by FSCJ

[For toxicological data from experimental animal studies]

I. Purpose

This guidance provides our basic stance and procedures for the use of the benchmark dose (BMD) approach for deriving a suitable point of departure (POD) in the risk assessment of chemicals in food with the aim to ensure greater consistency and transparency when the FSCJ applies the BMD approach in the risk assessment.

II. Definition

The definitions of terms in this guidance are as follows:

1. Benchmark Response (BMR)
   BMR is a change in the response level relative to the background response calculated from a fitted dose-response curve obtained by fitting a mathematical function (a mathematical model) to the overall dose-response relationship (i.e. the relationship between the level of exposure to a chemical and the frequency or magnitude of the effect caused by the given chemical).

2. Benchmark Dose (BMD)
   An exposure level of a chemical (i.e. a dose of chemical) corresponding to a predetermined BMR.

3. Benchmark Dose Approach (BMD approach)
   The approach derives a BMD and its confidence interval assuming that a specific mathematical function describes the dose-response relationship.

4. Benchmark Dose Lower Confidence Limit (BMDL)
   The lower confidence limit of the BMD.

Generally, the lower limit/bound of a two-sided 90% confidence interval or a one-sided 95% lower confidence interval on the BMD

5. AIC (Akaike Information Criterion)
   A Criterion for comparing the balance between model complexity and goodness-of-fit among different models
   AIC = −2 log (L) +2k
   log (L) : maximum log-likelihood of the model given the data
   k: number of estimated parameters included in the model

   The model providing a good fit with a few parameters has a relatively low AIC value, and is considered preferable.

6. Dichotomous data
   Quantal data, a type of categorical data, where an effect observed from exposure to a chemical in individuals, is classified into one of two possible outcomes, e.g. dead or alive, with or without tumor.

7. Ordinal categorical data
   A type of categorical data where an effect observed such as from exposure to a chemical in individuals to be allocated to several possible categories, where differences at different points of the scale are not necessarily equivalent, e.g. test results for urinary glucose level (−, ±, +, ++) and data for cancer stage.

8. Continuous data
   A type of data where effects observed such as from exposure to chemical measured on a continuum such as organ weight or enzyme concentration.

9. Restriction
   Restriction is required to limit the range of parameter(s) in the parameter estimation process. Restriction in the BMD approach limits range of parameters included in the mathematical model so that the shape of the dose-response curve derived from the mathematical model would be biologically plausible.

10. Bayesian estimation
    Parameter estimation of probability distribution based on Bayes' theorem.

III. Use of the BMD Approach in FSCJ's Risk assessment

This guidance provides basic guidance for the use of the BMD approach in the hazard characterization of chemicals to calculate the BMD and its confidence interval.

Expert judgment on biological significance as well as statistical analyses for calculating BMD and its confidence interval is required. Appendix "Concept for the Use of Bayesian Estimation on the BMD Approach" should be referred to when the Bayesian BMD approach is used to calculate the BMD and its credible interval.

In the case of adopting a concept or method not in line with this guideline, the details and reason shall be reported.

1. Collection and selection of dose-response data allowing the application of the BMD approach

(1) Collect all the available dose-response data such as observed data on specific doses and their corresponding responses for the chemical to be evaluated (hereinafter referred to as "relevant substance").

Depending on types and amount of data available, endpoints may be selected following the selection of toxicity studies for BMD analysis or the studies may be selected following the selection of endpoints under expert judgment.

(2) For the dataset selected using the procedures indicated in (1) above that is sufficient to establish a dose-response relationship, seek whether the following conditions apply in collaboration with experts.

A) The study design employed such as the selection of target animal species, number of animals, route(s) of administration, and dosage are appropriate.

B) The observed endpoint(s) has toxicological significance
Human extrapolation should also be considered.

C) Datasets examined for the feasibility of BMD analysis should include the following information.

a. Dichotomous data: the number of animals per dose group and number of animals showing response.
b. Ordinal categorical data: number of animals per dose group, number of animals showing response and degree of response in each individual (grades)
c. Continuous data: data per each individual
When individual data are unavailable, number of animals, central estimate (mean or median value) and a measure of response variability (standard deviation (SD) or interquartile range) for each dose group

It should be noted that when dataset is expected to include covariates (i.e. any variable that affects the outcome such as litter size in developmental toxicity study), the relevant information should be provided.

(3) To increase the number of samples to be analyzed and to increase confidence in the BMD analysis, datasets that are statistically and biologically compatible may be combined based on expert judgement. Statistical and biological judgement should be justified prior to combining the datasets to generate a reliable dose-response relationship. Characteristics of each study design including individual datasets affecting the dose-response pattern such as dose range, species, sample size, and data measurement method shall be explored. When available, it is recommended to use the original data (i.e. individual data) to construct the combined dataset since the original data would provide detailed information on the response in the dose region corresponding to the BMR.

(4) Careful consideration is needed when reducing (dichotomizing) the ordinal categorical data, data in which histopathological observations in treatment groups are reported by severity category (i.e. grade), to quantal data. Toxicological significance of the effects should be determined based on expert judgement taking into account types of histopathological observations, approach to categorizing the severity and incidence of observed effects. The incidence of observed effects in the control groups should also be examined. Consider the incidence of histopathological observations equal to or severer than the grade to be adverse, and treat the total incidence as quantal data.

2. Selection of BMR

(1) Dichotomous (Quantal) data

A) Calculations of BMD and its confidence interval (BMD calculation) require the selection of a BMR.

Considering the statistical power of data from experimental animal studies, comparing potencies across chemicals or endpoints and consistent[1] relationship between the BMD and the traditional NOAEL approach, an extra risk[2] of 10% selected for BMR is generally used for BMD calculation.

B) However, a BMR of 10% extra risk is not always suitable for all dose-response datasets. The BMR shall be selected considering the biological significance and the characteristics of the dataset to be used for modeling under expert judgement.

---

[1]S. Sand *et al*., A Signal-to-Noise Crossover Dose as the Point of Departure for Health Risk Assessment. Environmental Health Perspective. 2011, 119: 1766-1774.

[2]Extra risk: A measure of risk proportional to an increase in the risk of an effect adjusted for the background incidence of the same effect. An absolute change in frequency of response (additional risk, calculated as P (d)–P (0)) divided by the non-affected fraction in the controls.

Extra risk is calculated as follows: [P (d)–P (0)] / [1–P (0)].

P(d): The probability of response at a dose

P(0): The probability of response in the absence of exposure (i.e. background response in controls)

C) For example, when 10% extra risk (i.e. the response level) is much smaller than the response observed at the lowest dose, and the BMD calculation involves downward extrapolation to range much below the lowest dose, it may be desirable to use a BMR greater than 10% extra risk.

In a case where dose-response data relating to effects on F1 offspring from reproductive and developmental toxicity studies are used, it may be recommended to apply a BMR lower than 10% extra risk from a statistical standpoint.

(2) Continuous Data

A) When there is a scientific basis to establish a level of change in the continuous endpoint that is biologically significant, the defined level of change is used as a BMR.

B) A continuous data is expected to show various distributions, therefore, in the absence of biological significance in the level of change, select the BMR for each dose-response dataset in collaboration with experts.

C) When there is a scientific basis to establish a cutoff value of the endpoint, a hybrid[3] approach may be applied. After presuming a distribution of the data, specify the proportion of individuals who are expected to show higher/lower responses than the cutoff value in the control group. Then, a specified increase in the proportion of individuals who would show higher/lower responses than the cutoff value is defined as the BMR. In the Hybrid approach, calculate the dose corresponding to this defined BMR as the BMD.

D) When there is a scientific basis to establish a cutoff value of the endpoint and individual data are available, continuous data can be dichotomized based on the cutoff value to select the BMR. Although it should be noted that this approach results in a loss of information, in some cases it is desirable to calculate BMD with this defined BMR.

3. Dose-response modeling for BMD calculation and determining POD

(1) Dose-response modeling (model fitting)

A) All datasets that meet the conditions defined in 1. (2), above, should be considered for modeling to ensure that no endpoints with the potential to have the sensitive effects are excluded from the assessment.

B) Fit a model to all dose-response data that describes the dataset. Dropping data on specific dose group(s) such as the highest dose group should not be conducted unless there is a biological basis.

[3]K. S. Crump, Calculation of Benchmark Doses from Continuous Data. Risk Analysis, 1995. 15: 79-89 U.S. EPA, Benchmark Dose Technical Guidance. 2012.

C) Ideally, the dose-response relationship for a given chemical and endpoint would be described by a biologically based single mathematical model that describes toxicokinetics and toxicodynamics processes related to the chemical.

When such a model that describes an essential mechanism of action (hereinafter referred to as "biologically based model") is available, a priori selection of the model is recommended for modeling.

D) In the absence of a biologically based model, use a set of models implemented in BMD software or online tools (hereinafter referred to as "software") for modeling.

E) Use the software with well-documented experience and theoretical background that allows BMD calculation judged reasonable by experts.

F) If the model averaging approach is available in the software, also use the approach in the modeling process.

G) When model specific restriction can be activated on model parameters, use both a model with and without restriction(s) imposed on the parameter(s) for the modeling based on the assumption that two models are distinct and they produce different results.

(2) Evaluating the results of the dose-response modeling and BMD calculation

For each dataset, the result from fitting each of the models and model averaging shall be evaluated as follows.

A) Model averaging is used

Individual dose-response curves derived from fitted models used in model averaging is evaluated based on how well the model describes the dose-response relationship (e.g. assessing goodness-of-fit). Model averaging is adopted if the evaluation result is judged appropriate by experts.

Further, the model used for model averaging that does not provide an adequate description of the data (i.e. the resulting curve does not adequately fit the dose-response data, not supported by visual examination or goodness-of-fit[4]), may be rejected. In such case, remaining models that have met the criteria for adequacy may be used, while models with lack of fit are excluded.

B) Model averaging is judged unsuitable to provide adequate or reliable results or model averaging is not used

All the fitted models are evaluated as follows.

---

[4] Describes fit between the dose-response data and the dose-response curve resulting from dose-response modeling

a. Reject the resulting curve that does not provide an adequate description of the data, not supported by visual examination or goodness-of-fit statistic under expert judgement.
Use a value of 0.1 to compute the critical value for goodness-of-fit. If there is a priori reason to use a specific model more conventional, suitable values may be used.

b. All the fitted models are evaluated based on the following criteria from (a) to (d) based on expert judgement.
(a) The resulting curve is biologically plausible.
(b) The value of its AIC is within 3 units of the lowest AIC among all the fitted models.
(c) BMD estimates and its confidence intervals are sufficiently close with those of the remaining models.
(d) BMD estimates and the lower bounds of the confidence intervals are not significantly lower than the lowest dose observed.

c. Select all the dose-response curves obtained from a model that have met the criteria described in the above b. When no model fulfills the criteria, judge whether it is appropriate to apply the BMD approach to the dataset under expert judgement.

(3) Determination of POD

A single or multiple independent dose-response curves are obtained for a specific dataset from (2).
Select dose-response curves by identifying endpoints that may be relevance to human health considering the relevance of endpoints that supports human extrapolation, characteristics of the data and experimental design, as well as BMD and its lower confidence limit (BMDL) in collaboration with experts. The lower bound of 90% confidence interval of the BMD (BMDL) calculated from the selected curve will be used as a POD.

4. Reporting of the BMD analysis

In reporting a BMD analysis for a target substance, the following information should be documented to allow deliberation and assessment at the relevant Expert Committee.

(1) The software used, including the version number.

(2) Information on each dataset to which BMD analysis is applied.

A) The study design employed such as target animal species, endpoint, and route/duration of administration.

B) When open-access datasets are used, provide author(s) name, journal title, year of publication, and relevant information.

C) Information on dose-response data

a. Dichotomous (quantal) data: the total number of animals and the number of animals showing response for each dose level.

b. Ordinal categorical data: total number of animals, number of animals showing response and degree (grade) of response in each individual for each dose level.
When ordinal categorical data is reduced to quantal data, provide details of and rationale for the procedure employed.

c. Continuous data: number of animals, central estimate (mean or median value) and a measure of response variability (standard deviation (SD) or interquartile range) for each dose level.

D) When there is any variable considered as a covariate, provide the information.

E) When datasets are combined, the following should be provided: dataset information on C) and D), experimental design such as measurement of each dataset prior to combining the dataset, and justification for combining the datasets.

(3) Selection of BMR and rationale for its selection

(4) Result of dose-response modeling and BMD calculation to each dataset

A) Plot of fitted-dose-response curve including the dose-response curve derived from model averaging and unconstrained/constrained curves.

B) Information on each fitted model(s) including BMD with its 90% confidence intervals and model fit indices such as AICs.

C) Information on the statistical method used for the calculation of the confidence limit for the BMD (likelihood profile, bootstrapping methods and others).

D) When model averaging is used, provide weights used for each dataset and the weighting estimate for each model.

(5) Decision-making process for evaluating the fitted model for each dataset

(6) POD determined, the rationale for selection of the endpoint and dataset to derive the POD

IV. Re-evaluation

This guideline shall be reviewed when deemed necessary, referring to globally accepted risk assessment methods for the use of the BMD approach and the up-to-date use of the BMD approach in risk assessment.

Appendix

Provisional translation

Concept for the Use of Bayesian Estimation on the BMD Approach

1. Foreword

Given the advances in computational science, leveraging the Bayesian estimation to benchmark dose approach (hereinafter referred to as "the Bayesian BMD approach") is getting underway in recent years. This publically available tool for Bayesian BMD modeling is also facilitating the use in the regulatory context. Bayesian estimation has a technical advantage that improves estimation accuracy by using previous experimental data as a prior distribution. Environmental health criteria (EHC) monograph※1 and the EFSA guidance※2 described that Bayesian BMD approach is recommended as the preferred method to identify the point of departure for human health risk assessment. The FSCJ summarized important points to consider for using the Bayesian BMD approach in light of the global situation that the Bayesian framework is introducing to dose response assessment.

However, since some remaining issues need to be clarified for its use, the FSCJ considered that it would be difficult to promptly shift the paradigm of BMD approach from the existing frequentists to the Bayesian. This document is presented as the appendix in the revised FSCJ BMD guideline.

※1 The 2020 update to Chapter 5 of the Principles and Methods for the Risk Assessment of Chemicals in Food, Environmental Health Criteria 240, World Health Organization

※2 The Guidance on the Use of the Benchmark Dose Approach in Risk Assessment, European Food Safety Authority updated in 2022

2. The Bayesian BMD Approach

In frequentist estimates, the parameters are assumed to be fixed constants. According to the frequentist BMD approach, parameters of the mathematical model are inferred by maximum likelihood estimation[1], and then estimate the BMD using obtained dose-response model.

By contrast, the parameters are considered as a random variable in Bayesian estimate.
For procedure, at first, the prior distribution is considered as the probability distribution of parameters. Secondary, parameter values are estimated by relating its prior distribution and observed data to the posterior inference. Thus, in the Bayesian BMD approach, the prior

[1] Maximum likelihood estimation is an estimation method in which parameters are selected so as to maximize the probability of the observed data.

probability distribution of the model parameters are updated with obtained dose-response data, resulting in a posterior probability distribution for those parameters that the degree of uncertainty of those parameters, and then obtained the BMD as a probability distribution.

An external scientific report funded by an FSCJ research grant (JPCAFSC20202001※1) described the Bayesian BMD approach as having technical advantages, which include more stable estimation even when dealing with fairly small data, and more precise estimation of the criterion (e.g. BMDL) for POD by using to historical information to set informative prior.

※1 The 2020-2021 Research Program for Risk Assessment on Food Safety "Methodologies and global trends of using Bayesian estimation for bench-mark dose method."

3. Choice of the Prior Probability Distributions

In a Bayesian estimate, the prior probability distribution is updated using observed data to get the posterior probability distribution. In large samples, the effect of the prior probability distribution becomes small relative to the effect of the data. However, sample size of the available data used in risk assessment of food is often limited, and therefore the prior probability distribution may influence inference task of BMDL used as POD. Thus, choosing the prior should be carefully considered when applying the Bayesian BMD approach in the risk assessment of chemicals in foods.

Such prior distributions on the parameters can be employed the uninformative prior[2] and the informative prior[3]. A prior assumption on dose response functions should be prioritized using a flat uninformative prior when prior knowledge is either insufficient or inappropriate for use. An informative prior can utilize to improve the estimation accuracy, but may also lead to a bias on final outcome of the value of BMD.

When assigning a prior probability distribution, the following points should be considered;

- A prior probability distribution should be decided in each individual assessment based on expert knowledge.
- A sensitivity analysis, such as comparing the results obtained across from multiple choices for the prior distribution including the uninformative priors would be required when informative prior distributions are used to represent the prior information.

[2] Uninformative prior is a prior probability distribution formulated assuming there is no prior knowledge. For example, uniform distribution (in which probability of all outcome, i.e. densities, are equally likely within a specified range) is sometimes assigned to uninformative prior distribution. However, a narrow range uniform distribution may contain strong prior information, thus it is not necessarily assigned to uninformative prior distribution. Use of an uninformative prior is characterized by low arbitrariness and a tendency to yield outcomes similar to frequentist estimations.

[3] Informative prior is a prior probability distribution that relies on previous experimental data, literature, and expert knowledge. An informative prior based on scant knowledge is called weakly informative prior.

4. Model Averaging or Selection of Specific Models

In general, model averaging should be prioritized to address model uncertainties related to the variance over model estimate.

In case that model averaging is inappropriate, such as the obtained credible interval consider as too broad, the single specific model may be chosen from fitted dose-response model based on expert judge by referring a criterion of fit (e.g. the Bayesian Information Criterion (BIC)[4], and the highest posterior weight).

5. Caution Needed to Evaluating Results of Dose-response Modeling and BMD Calculation

As described in section 3, informative priors may enhance the reliability of BMD estimation but also lead to biased results. Therefore, when using informative prior instead of uninformative prior, the effect of employed prior should be well evaluated and presented by a comparison with the result of BMD estimation obtained from the modeling with multiple prior distributions including uninformative priors.

6. Reporting of the BMD Analysis

The rationale for the choice of the prior probability distribution should be explained and documented. For choosing an uninformative prior even when prior knowledge is available, the rationale for not choosing an informative prior must be documented. Likewise, when choosing an informative prior, the rationale for the choice and other related information should be stated to communicate that the choice is not arbitrary.

7. Going Forward

Trends and scientific findings, both domestic and abroad, will continue to be monitored with regard to future developments and new expertise pertaining to the Bayesian BMD approach. This Appendix will be updated accordingly.

[4] As with the Akaike information criterion (AIC), BIC is a criterion for comparing the balance between the mathematical model complexity and goodness-of-fit among different models with the measured data. Models with lower BICs are preferred.