Provisional translation

The decision of October 2019

Revised in April 2024

Food Safety Commission of Japan (FSCJ)

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

Guidance on the Use of the BMD Approach in Risk Assessment

Contents

Revision history

Date (edition)	Major revisions
October 2019 (First edition)	Guidance created as “For toxicological data from experimental animal studies”
June 2022 (Second edition)	Guidance modified based on results of the research project on the benchmark dose (BMD) approach.
September 2023 (Third edition)	Appendix “Concept for the Use of Bayesian Estimation on the BMD Approach” added.
April 2024 (Fourth edition)	New entry “Application of BMD to dose-response data obtained from epidemiological studies” added to guidance.

Section 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)[1] 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.

Section II Explanations of terms

The explanations of terms used in this guidance are as follows:

1. Benchmark response (BMR)

   BMR represents "a significant change in response level relative to the background response in terms of risk assessment."[2] The BMR is calculated from a fitted dose-response curve obtained by fitting a mathematical function (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. The BMR is used to calculate the benchmark dose.

2. Benchmark dose (BMD)

   BMD represents the exposure level of a chemical (i.e., a dose of the chemical) corresponding to a predetermined BMR.

3. Benchmark dose approach (BMD approach)

   The BMD approach derives a BMD and its confidence interval, assuming that a specific mathematical function (mathematical model) describes the 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.

4. Benchmark dose lower confidence limit (BMDL)

---

[1] Point of departure (POD): A value derived from the results of dose-response evaluations obtained from various animal studies and epidemiological studies, usually referring to the no-observed-adverse-effect level (NOAEL) or BMDL. It is used to set the health-based guidance value (HBGV) and to calculate the margin of exposure (MOE). Internationally, POD is also referred to as reference point. Source: Glossary of Food Safety Terms (Food Safety Commission of Japan; in Japanese language only)
http://www.fsc.go.jp/yougoshu/kensaku_hyouka.html

[2] "A significant change in response level relative to the background response in terms of risk assessment" includes the following examples, depending on the characteristics of the endpoint or the design of the study:
A) "The amount of change at a given exposure level" relative to "the value of the endpoint at zero or minimal exposure (baseline)"
B) "The amount of change at a given exposure level" relative to "the value of the endpoint at zero or minimal exposure (baseline)"

BMDL represents the lower confidence limit of the BMD. Generally, it is the lower bound of a two-sided 90% confidence interval or a one-sided 95% lower confidence interval on the BMD.

5. Akaike information criterion (AIC)

AIC is a criterion used to compare the balance between model complexity and goodness-of-fit among different models.

AIC = -2 log(L) +2k

log(L) : the maximum log-likelihood of the model given the data

k: the number of estimated parameters[3] included in the model

A model that provides a good fit with fewer parameters will have a relatively low AIC value and is considered preferable.

6. Dichotomous data

Dichotomous data, also known as quantal data, is a type of categorical data where an effect observed from exposure to a chemical in each individual is classified into one of two possible outcomes, e.g. dead or alive, presence or absence of tumor.

7. Ordinal categorical data

Ordinal categorical data is a type of categorical data where observed effects, such as those from exposure to a chemical, are allocated to several possible categories in each individual. The differences between different points on the scale are not necessarily equivalent, e.g., test results for urinary glucose levels (−, ±, +, ++) and data for cancer stages.

8. Continuous data

Continuous data refers to a type of data in which observed effects, such as those associated with exposure to a chemical, are measured on a continuum. Examples of such continuous measurements include organ weight and enzyme concentration. Discrete data, such as the number of tumors that have occurred, may sometimes be treated as continuous data.

9. Restriction

Restriction is required to limit the range of parameter(s) in the parameter estimation process. Restriction in the BMD approach limits the 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.

[3] If y＝f（x, a） represents the relationship between the response variable y and the explanatory variable x for various values of a, then a is called a parameter (even if there are multiple parameters in the model).

10. Bayesian estimation

Bayesian estimation refers to parameter estimation of probability distribution based on Bayes' theorem.

Section III Use of the BMD Approach in FSCJ's Risk Assessment

This guidance outlines the fundamental steps for utilizing the BMD approach in the hazard characterization of chemicals and focuses on calculating the BMD and deriving its associated confidence interval. Data collection, selection, and the adoption of optimal dose-response models should be conducted based on characteristics of data based on various data characteristics when employing the BMD approach, because there are differences in data features obtained from animal experiments and epidemiological studies (see sections I and II).

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 guidance, the details and the reasons as to why it was adopted should be reported.

I. Application of the BMD approach to dose-response data derived from experimental animal studies

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

(1) Collection of data from animal experiments

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 the types and amount of available data observing the dose-response relationships (hereinafter referred to as "dose-response data"), endpoints may be selected following the selection of toxicity studies for BMD analysis, or studies may be selected following the selection of endpoints. In either order, the process should be conducted under expert judgment..

(2) Selection of datasets

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 under expert judgement.

A) The study design adopted such as the selection of target animal species, number of each animals, route(s) of administration, and dose should be appropriate.

B) The observed endpoint(s) should have toxicological significance. Extrapolation to humans should also be considered, taking into account the results of epidemiological studies.

C) Datasets examined for the feasibility of BMD analysis should include the following information. Furthermore, upon dose-response modelling[4], it should be noted that when the dataset is expected to include covariates[5] (i.e., any variable that affects the outcome such as litter size in the developmental toxicity study), relevant information should be provided.

a. Dichotomous data: 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 the grades of their response[6].

c. Continuous data: data per each individual. When individual data are unavailable, the number of animals, the central estimate (mean or median value) and the measure of response variability (standard deviation (SD) or interquartile range) for each dose group should be provided.

(3) Acquiring the original data

Utilizing original data (i.e., individual data) provides detailed information on the response within the dose region corresponding to the BMR. Consequently, it is recommended to refer back to the original data and utilize them whenever feasible.

(4) Combining datasets

To increase the number of samples available and to increase confidence in the BMD analysis, datasets that show homogeneous characteristics could be combined under expert judgement. In these cases, characteristics of each study design for the individual datasets such as dose region, species, sample size, and data measurement methods should be explored to prevent ambiguous dose-response relationships. Furthermore, the validity of the combination should be examined from statistical and biological perspectives prior to combining the datasets.

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

---

[4] Deriving a fitted dose-response curve obtained by fitting a mathematical function (mathematical model) to the dose-response relationship, i.e., the relationship between the level of exposure to a chemical and the effect caused by the given chemical.

[5] Variables that are not the primary object of interest but are considered to influence the response variable and added for analysis.

[6] The degree of ordered effects observed (pathological findings, etc.). Careful consideration is necessary when reducing ordinal categorical data, in which incidences are reported by severity category (i.e., grade) to dichotomous data. Specifically, considering the incidence of observed effects in the control groups, the presence or absence of histopathological observations graded higher than those judged significant from a toxicological perspective—based on type, grading method, and incidence—could be used as dichotomous data under expert judgement.

potencies across chemicals or endpoints and consistent[7] relationship between the BMD and the traditional NOAEL approach, an extra risk[8] 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 dose-response data to be used for modeling, under expert judgement. 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 a range much below the lowest dose, it may be desirable to use a BMR greater than 10% extra risk. In the 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.

(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 dose-response data is expected to show various distributions; therefore, in the absence of biological significance in the level of change, the BMR is selected for each dose-response data under expert judgement.

C) If there is a scientific basis to establish a cutoff value for the endpoint, alternative methods such as the hybrid approach[9] may be adopted. The hybrid approach first presumes a distribution of the data, whereby the proportion of individuals who are expected to show higher/lower responses than the cutoff value relative to the control group is specified, after which a specified increase in the proportion of individuals who would show higher/lower responses than the cutoff value is predefined as the BMR, based on which the dose corresponding to this BMR is calculated and defined as the BMD.

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

---

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

[8]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)

[9] Details on the hybrid approach has been suggested in the following: K. S. Crump, Calculation of Benchmark Doses from Continuous Data, Risk Analysis, 1995, 15(1): 79-89; U.S. EPA, Benchmark Dose Technical Guidance, 2012.

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.

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 software that has a robust track record, is well-documented, and possesses a theoretical foundation. The software must enable BMD calculations under expert judgement..

F) If the model averaging approach[10] 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 results from model averaging and dose-response modelling by fitting each of the models shall be evaluated as follows.

A) For cases involving model averaging

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.

---

[10] A method to calculate BMD-related indices by using mathematical models that fit the data well. It involves deriving a new model from a dose-specific weighted average of estimated responses, with weights determined by indices such as AIC.

Further, if the model employed for model averaging fails to adequately describe the data (i.e., if the resulting curve does not fit the dose-response data well, and lacks support from visual examination or goodness-of-fit[11]), it may be deemed unsuitable and rejected. In such cases, remaining models that have met the criteria for adequacy may be used, while models with lack of fit are excluded.

B) For scenarios where model averaging is considered inadequate for producing reliable results or when model averaging is not employed

All the fitted models are evaluated as follows.

a. Reject the resulting curve that is not considered to adequately describe the dose-response data based on visual examination or goodness-of-fit statistics 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 specific models that are more conventional, suitable values may be used.

b. All the fitted models are evaluated based on the following criteria from (a) to (d) based under 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 models that meet 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

Single or multiple independent dose-response curves are obtained for a specific dataset as a result of the evaluation of (2).

Select dose-response curves considering the significance of specific endpoints, etc., that support human extrapolation, characteristics of the range of dose-response data including the range of the exposure levels observed, and BMD and its lower confidence limit (BMDL) in collaboration with experts. The lower bound of the 90% confidence interval of the BMD (BMDL) calculated from the selected curve will be used as a POD.

---

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

4. Reporting of the BMD analysis

In reporting a BMD analysis for the relevant 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) Summary of study design adopted 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 any variables are considered as covariates, information should be provided.

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

(3) Selection of BMR and rationale for its selection

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

A) Plot of fitted-dose-response curve including the dose-response curve derived from model averaging and both restriction ON/OFF condition.

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

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

D) When model averaging is used, provide indices used for weighting and weights used for each model.

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

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

II. Application of BMD to dose-response data from epidemiological studies

1. Collection and selection of dose-response data to which the BMD approach is applied

(1) Collection of results from epidemiological studies

Systematically collect all relevant results from human epidemiological studies that investigate exposure levels of the relevant substance and observe its association with adverse effects or their indicators (hereinafter referred to as "health effects, etc.").This collection and selection process should consider health effects, etc., for which a causal relationship has been identified or suggested, based on the results of animal experiments and mechanisms, etc., in a comprehensive manner involving experts

(2) Selection of datasets

Of all results of the epidemiological studies collected in (1) above, select appropriate dataset(s) that can be judged reliable and valid for obtaining the POD, involving experts after organizing the following items and considering the application of the BMD approach.

Statistical significance of the dose-response relationship will not be considered.

A) The study design
   a. Classification (observational study, intervention study, etc.)
   b. Study period
   c. Follow-up and retrospective observation period, etc. (if applicable)

B) Information on the study population
   a. Characteristics of the study population (gender, age, race, etc.)
   b. Number of the study subjects
   c. Study area(s)

C) Information on exposure to the relevant substance
   a. Cause of exposure (through daily life, occupational, accidents, etc.)
   b. Route of exposure
   c. Duration of exposure
   d. The methods used to estimate exposure levels, including exposure indices and the measurement techniques employed
   e. Range of exposure levels
   f. If applicable, the exposure classification boundary (equal or logarithmic equal, interval, or quantile, etc.) and the number of individual study subjects in each exposure classification, etc.

D) Target health effects, etc., and their selection criteria (clinical significance, etc.)

E) Confounding factors considered in the study and whether and how they were adjusted

F) Approaches to mitigate biases and relevant considerations

(3) Acquiring original data

Using original data (individual study subject data), allows for a precise analysis of health effects, etc., at exposure levels corresponding to the BMR. Therefore, when feasible, it is advisable to obtain and utilize such original data or other relevant information.

(4) Combining datasets

When multiple datasets are available, they can be combined and analyzed after considering the dose-response relationship involving experts. This approach extends the exposure range and increases the sample size, thereby enhancing confidence in the BMD analysis. In such cases, the datasets should be combined after assessing their homogeneity based on the range of exposure, study subject characteristics, sample size, and data measurements. Also, justification of clinical and statistical significance is necessary to prevent an unreliable dose-response relationship.

2. Selection of BMR

When applying BMD to dose-response data obtained from epidemiological studies, the BMR should be selected based on toxicological, clinical, and public health considerations under expert judgement.

The magnitude (severity, duration, and incidence, etc.) and precision (or variability) of health effects, etc., should be considered in the selection. It should be noted that even in human epidemiological studies, BMD values calculated from the chosen BMR do not directly correspond to health-based guidance values (hereinafter referred to as “HBGV.”)

(1) Dichotomous (Quantal) data

In “I. Application of the BMD approach to dose-response data derived from experimental animal studies” in this guidance, “an extra risk of 10% selected for BMR is generally used for BMD calculation” in dichotomous data. On the other hand, setting a default BMR is challenging in epidemiological studies compared to animal studies due to variations in study subjects, study designs, and other factors. Therefore, when applying the BMD approach to dichotomous (quantal) data from epidemiological studies, the appropriate BMR should be determined for each specific dataset, health effect, etc., under expert judgement.

(2) Continuous data

A) If a significant change in laboratory values, etc., that is relevant at the population level through public health perspective can be set based on scientific evidence, the amount of change should be set as the BMR.

B) If a significant change in laboratory values, etc., relevant to public health perspective cannot be set under expert judgements, BMRs are set based on the association between the true outcome and laboratory values, etc., as well as background response rates (response levels) and their variability.

C) Apart from the situations described in the above scenarios A) and B), if there is a scientific basis to set a cutoff value for the endpoint, alternative methods such as the hybrid approach may be adopted. The hybrid approach first presumes a distribution of data, whereby the proportion of individuals who are expected to show higher/lower responses than the cutoff value relative to background is specified, after which a specified increase in the proportion of individuals who would show higher/lower responses than the cutoff value is predefined as the BMR, based on which the dose corresponding to this BMR is calculated and defined as the BMD.

D) When there is a scientific basis to set a cutoff value for an endpoint, and when individual study subject data is available, continuous data can be dichotomized based on the cutoff value to help select the BMR. However, it should be noted that this approach results in some loss of information. Despite this, in some cases, it is desirable to calculate BMD-related indicators using 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 sensitive effects are excluded from the assessment.

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

C) When covariates useful for calculating the BMD etc. are available, their potential usage should be considered under expert judgement.

D) Ideally, the dose-response relationship for a given chemical and its health effects, etc., 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 (the “biologically based model” as mentioned previously) is available, a priori selection of the model is recommended for modeling.

E) In the absence of a biologically based model, use a set of models implemented in BMD software or online tools (“software” as mentioned previously) for dose-response modeling.

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

G) If the model averaging approach is available in the software, also use the approach in the dose-response modeling process.

H) When restrictions can be activated to model parameters, use both a model with and without restrictions (i.e. restrictions ON/OFF) imposed on the parameters for the dose-response modeling. This approach, which uses individual models, is based on the assumption that two models are distinct and produce different results.

(2) Evaluating the results of the dose-response modeling

For each dataset, the results from model averaging and dose-response modelling by fitting each of the mathematical models shall be evaluated according to the following classifications.

A) For cases involving model averaging

Dose-response curves derived from fitted individual mathematical models used in model averaging are 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, any model used for model averaging that does not adequately describe the data (i.e., the resulting curve does not fit the dose-response data well, either visually or by goodness-of-fit), may be rejected. In such cases, remaining models that have met the criteria for adequacy may be used, while models with lack of fit are excluded.

B) Cases in which model averaging is considered inadequate for producing reliable results or when model averaging is not employed

All the fitted models are evaluated as follows.

a. Reject any resulting curve that is not considered to adequately describe the dose-response data, either visually or based on goodness-of-fit, under expert judgement. Use a statistical significance level of 0.1 to test goodness-of-fit. If there is a priori reason to use a specific, more conventional model, a suitable significance level 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 shape of dose-response curve is biologically plausible

(b) The difference between its AIC and the lowest AIC among all the fitted models is within +3

(c) The confidence intervals of BMD estimates are not sufficiently wider than the remaining models

c. Select the results obtained from dose-response models that meet all criteria described in the above “b.” When no model fulfills the criteria, determine whether it is appropriate to apply the BMD approach to the dataset under expert judgement

(3) Determination of POD

Results from a single or multiple independent dose-response modelling are obtained for a every dataset as a result of the evaluation of (2).

Select dose-response modelling considering the significance of specific health effects, etc., characteristics of the dose-response data including the range of the exposure levels observed, and BMD and its lower confidence limit (BMDL) under expert judgement. The lower bound of

the 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, the following information should be documented for use in the deliberation and assessment by the relevant Expert Committee for the relevant substance.

(1) The software used, including the version number

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

A) Data source (author names, journal titles, year of publication, name of organization managing the original data, etc.)

B) Summary of study design (information on the study population, health effects, etc., and route/duration of exposure, etc.)

C) Summary of research findings

a. Information on exposure to the relevant substance

b. Number of study subjects in each exposure category

c. Information on response level in each exposure category

d. Variability of "c." above

D) When information including health effects, etc., are converted (dichotomized, etc.,) details on the process and their rationale

E) Information when there is any variable considered as a covariate

F) When datasets are combined, the following should be provided: dataset information on C) and D), study design in each dataset, such as measurement method, and justification for combining the datasets

(3) Selection of BMR and rationale for its selection

(4) Result of dose-response modeling when BMD method is applied for each dataset

A) Plot of results from each dose-response modelling including the dose-response curve derived from model averaging and both restriction ON/OFF condition

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

C) Information on the statistical method used for the calculation of the confidence interval for the BMD (e.g. likelihood profile, bootstrapping methods)

D) When model averaging is used, provide indices used for weighting and weights used for each model.

(5) The process for evaluating the results from the dose-response model for each dataset

(6) The value of POD and the rationale for selecting each health effect and dataset used to derive the POD, etc.

Section IV. Re-evaluation

This guidance shall be reviewed when deemed necessary, referring to globally accepted risk assessment methods and the latest applications 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.