Provisional translation
Decision of February 2021
Food Safety Commission of Japan (FSCJ)

Guideline for Assessing Mutagenicity Using (Q)SAR in Risk Assessment

Chapter 1: Background

(Q)SAR ((Quantitative) Structure-Activity Relationship) is an *in silico* assessment method for inferring the toxicity of a chemical substance based on a database of *in vivo* and *in vitro* toxicity test data for various chemical substances. (Q)SAR is a method for estimating the effects and toxicities of a chemical substance based on the relationship between its chemical structure and its biological activity.

The accuracy of (Q)SAR predictions has been improving in line with the availability of more comprehensive toxicity databases. Accompanying this improvement in prediction accuracy, numerous software packages targeting various toxicities have been released as support tools (hereinafter referred to as (Q)SAR tool(s)), and they are being used for assessment of toxicities in the pharmaceutical and other sectors.

As a consequence of the supportive role of (Q)SAR in toxicity assessment, the accuracy of expert judgment has been enhanced and as such, increased robustness of assessment can be anticipated. Consequently, based on the recommendation of the Assessment Methodology Development Working Group, which states that it is highly significant to work toward the increased utilization of (Q)SAR[1], the Food Safety Commission of Japan has been accumulating scientific knowledge, including comparison and validation data based on toxicity test data, in order to utilize (Q)SAR in Risk Assessment.

Based on the scientific knowledge accumulated so far and to facilitate utilization of the methodology, a standard procedure for the use of (Q)SAR prediction results in Risk Assessment was established.

[1] Report from the Assessment Methodology Development Working Group, Food Safety Commission of Japan; Consideration points for Assessment Techniques for a New Era—Utilization of (Q) SAR and Read Across for Assessing the Toxicity of Chemical Substances —. 2017

Chapter 2: Basic Concept

At present, (Q)SAR is capable of toxicity prediction for chemical substances without conducting conventional toxicity tests, although it does not completely replace them. Therefore, it is assumed that (Q)SAR results can be used as a substitute for toxicity test data for toxicity assessment of chemical substances for which it is difficult to obtain toxicity test data. Such substances include substances eluted from food utensils and containers/packaging, impurities generated in the manufacturing process of chemical substances for which Risk Assessment is conducted (hereinafter referred to as a target substance), and metabolites of the target substance.

Even if toxicity test data are available for the target substance (Q)SAR results can be used to support expert judgment in cases in which data are limited, or the results differ among toxicity tests. However, (Q)SAR technology is still under development for some toxicity endpoints, and the prediction accuracy varies for each of these endpoints. Therefore, based on the status of the toxicity databases, (Q)SAR tool(s) and prediction accuracy[2], the standard procedure for assessing mutagenicity detected by the Ames test using the prediction results from (Q)SAR will be defined at first. Hereinafter in this guidance, mutagenicity detected by the Ames test is referred to as “mutagenicity”.

In using (Q)SAR in Risk Assessment, the nature of the target substance to be assessed, the purpose of the assessment, and available information on mutagenicity tests should be considered. Further, expert judgment in biology, physical chemistry, toxicology, chemo informatics, and other scientific fields necessary for the prediction of mutagenicity by (Q)SAR and for the judgment of prediction results, should be based on the opinions of experts in the respective fields.

[2] Benigni R, Battistelli CL, Bossa C, Giuliani A, Fioravanzo E, Bassan A et al. Evaluation of the applicability of existing (Q) SAR models for predicting the genotoxicity of pesticides and similarity analysis related with genotoxicity of pesticides for facilitating of grouping and read across. EFSA Supporting Publications 2019; 16: 1598E

Chapter 3: Mutagenicity Assessment Using (Q)SAR

In Risk Assessment, the standard procedure for assessing the mutagenicity of a target substance using (Q)SAR prediction results is as follows (see reference 1 for an overview of the procedure).

1. Preparation of target substance data

(1) For each target substance, confirm its name and prepare the following information. Multiple data sources should be referred to, and it should also be verified that there are no discrepancies among the sources.

a) Name
b) Regulatory numerical identifiers (CAS No., etc.)
c) Molecular formula
d) Structures (SMILES, etc.)
e) Molecular weight
f) logPow (While experimental values can be used, calculated values are normally used, and information on the software used in the calculations should also be provided.)

(2) If any of the following applies to the target substance identified in (1), they are not applicable to this guidance. Even if they are not included in the following, there are some chemical substances that cannot be predicted using (Q)SAR tools, including organometallic substances that are generally considered to be unsuitable for toxicity prediction by this method. Accordingly, when selecting the (Q)SAR tool as described in 2. below, applicability of the prediction model should be verified.

a) When there is no information on the target structure available or when the information available is deemed unreliable.
b) Any of the following substances
   i) Inorganic chemical compounds
   ii) Polymers
   iii) Proteins and other biological components
   iv) Mixtures[3]

2. Selection of (Q)SAR tools

To improve prediction accuracy, select one of the available (Q)SAR tools that performs

[3] If no components fall under any of a) and b) i)-iii), the toxicity of each component can be predicted individually and the toxicity of the mixture as a whole estimated from the results.

expert rule-based predictions and one that performs statistical-based predictions from among those that are judged to satisfy the following conditions[4,5].

a) The sensitivity and specificity of the model are indicated, and the scientific validity of the model can be verified in accordance with the OECD Principles for the Validation, for Regulatory Purposes, of (Quantitative) Structure-Activity Relationship Models[6];
b) The output format of the report is in accordance with the (Q)SAR Prediction Reporting Format (hereinafter referred to as "QPRF")[7], and the following information including the prediction results can be outputted:
    i) Name of the target substance;
    ii) Data regarding chemical structure of the substance (molecular formula, structural formula, etc.);
    iii) Data regarding physicochemical properties of the substance (logPow etc.);
    iv) Names and versions of the (Q)SAR tools and the prediction models used;
    v) Information on the reliability of the mutagenicity predicted by the (Q)SAR tools;
    vi) Information which support the predicted results of mutagenicity.

3. Mutagenicity predictions using (Q)SAR tools

(1) Target substance is analyzed to predict mutagenicity using the latest version of each (Q)SAR tool selected in 2. that is available at the time when the prediction is made. In the process, if necessary, the chemical structure information of the target substance is optimized to be suitable for input to the (Q)SAR tools by removing salt and hydration water, etc.

(2) For each (Q)SAR tool, output a report including mutagenicity prediction results and the data from 2. (b) and furthermore, compile it together with the following information to be used in provisional judgment of mutagenicity based on the prediction results by the (Q)SAR tool as outlined in 4. and assessment of mutagenicity as described in 5.

---

[4] Food Safety Commission of Japan. Research Program for Risk Assessment Study on Food Safety: Study on application of in silico evaluation method to risk assessment of trace level chemicals unintentionally contained in food, 2020.

[5] International Conference on Harmonization of Technical Requirements for the Regulation of Pharmaceuticals (ICH) M7 Guideline: Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk (Notification of the Director, Assessment and Management Division, Pharmaceuticals and Consumer Health Bureau, Ministry of Health, Labour and Welfare, No. 1110-3, dated November 10, 2015)

[6] Principles for validating (Q)SAR models used for regulatory purposes as indicated by the OECD. Five principles are listed: A defined endpoint, 2) an unambiguous algorithm, 3) a defined domain of applicability, 4) appropriate measures of goodness-of-fit, robustness and predictively, and 5) a mechanistic interpretation, if possible.

[7] NAFTA. (Quantitative) Structure Activity Relationship [(Q)SAR] Guidance Document, 2012.

a) Expert rule-based (Q)SAR tool prediction
   i) The test data if the (Q)SAR tool presents data from an Ames test on the target substance;
   ii) Molecular mechanisms of mutagenicity postulated from structural alerts and supporting test data;
   iii) Mutagenicity information for alert structures of target substances presented by the (Q)SAR tool if indicated;
   iv) If the (Q)SAR tool does not present an alert structure, then mutagenicity data for structurally similar chemical substances or partial chemical structures to the target substance if indicated;
   v) Information provided by the (Q)SAR tool itself regarding the reliability of its predictions

b) Statistical-based (Q)SAR tool prediction
   i) The test data if the (Q)SAR tool presents data from an Ames test on the target substance;
   ii) Degree to which the target substance is included in the applicability domain of the prediction model.
   iii) Mutagenicity data for structurally similar chemical substances with the target substance if indicated;
   iv) Test data that forms the basis of the mutagenicity predicted by the (Q)SAR tool;
   v) Information provided by the (Q)SAR tool itself regarding the reliability of its predictions

(3) If the prediction results of the (Q)SAR tools correspond to any of the following, a new (Q)SAR tool that satisfies the conditions described in 2. can be selected and used to predict the mutagenicity.

a) The target substance is out of domain of the (Q)SAR tool;
b) If the (Q)SAR tool determines that the target substance contains an unclassified or incalculable structure.

4. Provisional judgment of mutagenicity based on (Q)SAR tool prediction results

(1) Carry out expert rule-based and statistical-based (Q)SAR tool predictions

Prepare the table that classifies the prediction results obtained each (Q)SAR tools into "positive," "negative," or "unpredictable" (hereinafter referred to as "Summary Table").

(2) Classify the prediction results of the target substance obtained from each (Q)SAR tool as

"positive," "negative," or "unpredictable" based on the information from 3. (2) and the information in the Summary Table referred to in (1). When classifying the results as "positive" or "negative," the reliability of the prediction results is also classified as "high" or "low."

(3) Based on the classification results of (2), a provisional judgment of mutagenicity is made as follows.

a) Both (Q)SAR tool predictions are classified as negative:

i) If the reliability of both prediction results is classified as "high," the result is judged as "negative;"

ii) If the reliability of both or either of the prediction results is classified as "low," an additional prediction of mutagenicity using a new (Q)SAR tool that satisfies the conditions of 2. is performed, and the result of the prediction and its reliability are also considered to assess whether the result is judged as "negative" or not. If it is difficult to determine whether a negative result is possible, then the result is judged to be "indeterminable."

b) Both (Q)SAR tool predictions are classified as positive:

i) A "positive" result is obtained unless the reliability of both prediction results is classified as "low;"

ii) When the reliability of both prediction results is classified as "low," they are comprehensively reviewed by cross-checking the information on which the positive judgment was based, and if the mutagenicity of the target substance is assessed to be positive, the result is judged as "positive." Otherwise, it is assumed to be "indeterminable." Further, in Risk Assessment, if the mutagenicity of the target substance needs to be assessed solely by the prediction results of the (Q)SAR tools, it is judged as "positive."

c) Conflicting classification of prediction results by (Q)SAR tools

i) If the reliability of the positive prediction is classified as "high," it is judged as "positive;"

ii) If the reliability of the positive prediction is classified as "low," the result is considered "indeterminable."

d) Either or both (Q)SAR tool predictions are classified as unpredictable

i) If both predictions are "unpredictable." the result is "indeterminable;"

ii) If a prediction by one tool is unpredictable and the prediction by the other tool is positive, and if the reliability of the other is classified as "high," then it is judged

as “positive;”

iii) If a prediction by one tool is unpredictable and the other tool prediction result is positive, and moreover if its reliability is classified as being low, data that forms the basis of the judgment is comprehensively reviewed and if the mutagenicity of the target substance is assessed to be positive, the result is judged as “positive.” Otherwise, it is assumed to be “indeterminable.” Further, in Risk Assessment, if the mutagenicity of the target substance needs to be assessed solely by the prediction results of the (Q)SAR tool, it is judged as “positive;”

iv) If one of the predictions is unpredictable and the other is negative, the result is considered “indeterminable.”

5. Mutagenicity assessment

In addition to the provisional judgments based on the (Q)SAR tool predictions made in 4. and the data collated in 3. (2), Ames test results for the target substance as well as for substances with similar chemical structures to the target substance and other data that can be utilized to determine the mutagenicity of the target substance are taken into consideration to undertake a final assessment of mutagenicity.

6. Description of mutagenicity assessment results in report

An assessment report on the mutagenicity of the target substance based on (Q)SAR should include the final assessment of the mutagenicity determined in 5. as well as its basis. In addition, in order to ensure transparency in the Risk Assessment, the following information on the analyses performed should also be included for each individual (Q)SAR tool used.

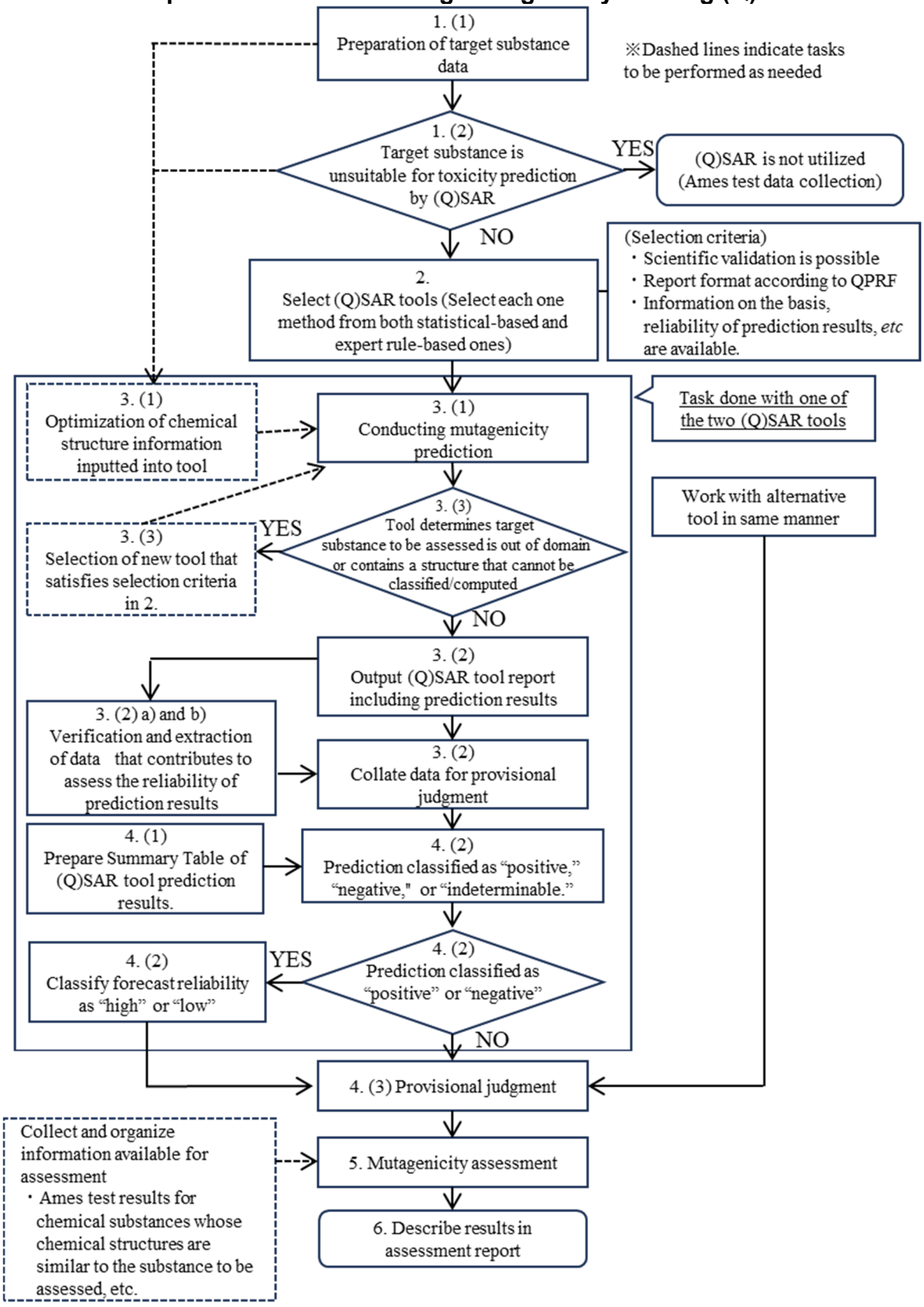
a) Name of the Target substance;
b) Structural formula used in the prediction based on the (Q)SAR;
c) Name(s) of the (Q)SAR tool and prediction model and version;
d) Classification of the prediction result from each (Q)SAR tool in 4. (2) and classification of its reliability.

Chapter 4: Review of the Implementation Procedure

This guidance will be revised as necessary based on the domestic and international trends of (Q)SAR and the results of their utilization in Risk Assessment.

(Reference 1)

The procedure for assessing mutagenicity utilizing (Q)SAR



(Reference 2)

Terminology

For general technical terms used in this guidance, refer to the latest "Glossary of Food Safety Terms[8] prepared by the Food Safety Commission of Japan. The explanation of terms not included in the Glossary is as follows.

1. SMILES (Simplified Molecular Input Line Entry System)

   A method for standardizing the chemical structure of molecules by means of computer-friendly, two-dimensional descriptors. That is, each non-hydrogen atom (hydrogen is included only in special cases) is indicated by its symbol. Double bonds are indicated by "=," triple bonds by "#," branches by parentheses, and rings by numbers. For example, CCO means ethanol, CC(=O) O means acetic acid, C#N means hydrogen cyanide, and c1ccccc1N means aniline (the numbers are the starting and ending points of the ring, and the lower case "c" means aromatic carbon).

2. logPow (octanol/water partition coefficient)

   The logarithm of the ratio of a compound's concentration in octanol to its concentration in water at equilibrium for a substance dissolved in a mixed solvent of octanol and water. A high value implies that the substance is highly hydrophobic.

3. Prediction model

   A theoretical equation, algorithm, or program used to predict the action or toxicity of a chemical substance from measured values obtained by experiment or other means. Expert rule-based and statistical-based prediction models are commonly used in (Q)SAR, and this guidance deals only with these two types.

4. Expert rule-based

   A type of (Q)SAR prediction model. From known data, characteristic substructures (alert structures) of chemical substances that cause positive results in the Ames test are defined, and the results of the Ames test are qualitatively predicted based on empirical rules.

5. Statistical-based

   A type of (Q)SAR prediction model. The structures of chemical substances and their physical properties are transformed into descriptors, and descriptors that are highly correlated with the Ames test results are used to predict test results. Descriptors include the presence or absence of fragments in chemical substances and numerical values that describe

[8] http://www.fsc.go.jp/yougoshu.html

the electronic and physicochemical properties of the molecules.

6. Applicability domain

   The range within which a prediction model is expected to produce reliable prediction results. It is usually defined by the structural characteristics of the substance in the measured data set used to create the prediction model, or by the range of descriptors.