Evaluation report on the cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests

Yoji Takahashi[1], Toshihiko Kasahara[2], Shinobu Wakuri[3], Teruaki Kodaira[4], Hisakazu Iwai[5], Shigenobu Hagino[6], Shigeo Kono[7], Yomiko Iwase[8]

[1]National Institute of Health Sciences, [2]FUJIFILM Corporation, [3]Hatano Research Institute, Food and Drug Safety Center, [4]Asahi Kasei Pharma Corporation,

[5]Sanwa Chemistry Co., Ltd., [6]Shiseido Co., Ltd., [7]CMIC Bioresearch, [8]Mitsubishi Tanabe Pharma Corporation

Summary

This report is a peer review by the JaCVAM Peer Review Panel on alternatives to acute toxicity testing, which reviewed background review documentation (BRD) prepared from an ICCVAM peer review of a validation study on *in vitro* Cytotoxicity Test Methods for Estimating Starting Doses Used in Acute Oral Systemic Toxicity Tests.

Acute toxicity tests evaluate potential hazards by chemical substances using rodent cell lines. The LD_{50} (lethal dose, producing lethality in 50% of test animals) values obtained from acute toxicity testing are used as criteria for judgments based on the Globally Harmonized System of Classification and Labeling of Chemicals (GHS) as well as those based on Japan's Poisonous and Deleterious Substances Control Act. Criticism has been voiced, however, about test methods that use the death of animals as an endpoint, and there is also debate over the usefulness of acute toxicity test data as a reference for excessive ingestion of chemical compounds by humans. Alternative methods utilizing cytotoxicity testing have been studied, but there are no tests at present that have been evaluated by regulatory agencies as acceptable for reliability, validity, usefulness, and range of application. At present, acute oral systemic toxicity tests available under OECD toxicity test guidelines include the Fixed Dose Procedure (OECD 420:FDP), Acute Toxic Class method (OECD 423:ATC), and Up-and-Down Procedure (OECD 425:UDP). These test guidelines are designed to determine an LD_{50} value for a chemical compound through administration of a predetermined dose from a regime of four or eight total dose levels, with successive tests at a lower dose if mortality occurs or at a higher dose if mortality does not occur. Accordingly, selection of a suitable starting dose is the key to reducing the total number of animals needed for testing.

This test method is designed to reduce the total number of animals need for testing by means of an in vitro approach for predicting starting doses for acute toxicity tests based on a cytotoxicity test. Specifically, a neutral red uptake (NRU) cytotoxicity test is used to determine a value for IC_{50} (mM), after which an initial dose for the acute toxicity test is estimated using a regression model to compare values of LD_{50} (mmol/kg) and IC_{50}, based on RTECS® data. Validation of this test method was performed using a cytotoxicity test for 72 reference compounds chosen to form a representative selection distributed across the entire range of GHS classification for acute oral toxicity as well as a human cell line of normal human epidermal keratinocytes (NHK) and a rodent cell line of BALB/c mouse fibroblasts (3T3).

Although a regression model for predicting starting doses for acute toxicity tests has been established based on correlation of values for LD_{50} and IC_{50}, we do not yet have a sufficient understanding of the mechanism that equates cell death with animal death. Correlation is not necessarily indicative of a cause-effect relationship. Also, compounds that are soluble, precipitative, or volatile as well as those exhibiting metabolic toxicity or specific toxicity toward the liver, central nervous system, kidneys, heart, lungs, or hematopoietic organs are to be excluded from this test, because they cannot be evaluated by means of a cytotoxicity test.

Although the cytotoxicity test method is considered a suitable approach, it is necessary to afford flexibility in terms of dosages, media, *in vitro* equipment used to prepare test substances, and methods for impurity doping of test substances, provided test accuracy can be maintained.

Evaluating the predicted values for LD_{50} obtained by this test method across the entire range of GHS classification, we find predictivities of 31% (21/67) for 3T3 and 29% (20/68) for NHK cells, with particularly low predictivity for virulent compounds. In terms of mechanisms of toxicity, outliers were found for compounds that exhibit effects on the central nervous system and heart. These results, too, show low predictivity for compounds that exhibit organ-specific toxicity, indicating a lack of suitability for evaluation.

Validation was performed at three facilities, either in accordance with or in the spirit of GLP. We further determined that there were no issues with the quality of the data.

Reduction of the total number of animals needed was evaluated using computer simulation. Results indicated reductions of between 0.49 and 0.66 animals for UDP and between 0.51 and 1.09 animals for ATC. In particular, for substances with low toxicity of $LD_{50} > 2000$ mg/kg or > 5000 mg/kg, reductions of between 1.28 and 1.65 animals for UDP and between 2.03 and 3.33 animals for ATC were indicated. The question of whether the use of this test will actually reduce the total number of animals needed for testing cannot be answered without further study that compares results of compound tests for which the total number of animals used is included in the test data with predictions of the total number of animals needed based on starting doses predicted by this test. Also, if this test method is introduced into test guidelines, it will be necessary to perform an aggregate analysis of test data to verify whether or not reductions in the total number of animals used were actually achieved. It is further noted that low predictivity for virulent chemical substances will not contribute to reducing suffering inflicted upon animals.

Based on the above, we find that ICCVAM's peer review of a validation study of *in vitro* Cytotoxicity Test Methods for Estimating Starting Doses Used in Acute Oral Systemic Toxicity Tests addresses all items, processes, and data necessary to a validation, and see no problem in accepting the findings of the ICCVAM validation. Since this test method demonstrates predictivity for compounds with low toxicity and has the potential to help reduce the total number of animals needed in testing, we find it can be used as necessary to inform estimates of starting doses for acute toxicity tests. Low predictivity for compounds classified as virulent, however, suggests that this test has little to contribute in terms of reducing the total number of animals needed in testing or reducing the suffering inflicted upon animals. This test method, therefore, is not suitable for evaluating compounds that exhibit organ-specific toxicity and is difficult to implement for substances that exhibit volatility or low solubility. We therefore find that it would be unreasonable to make blanket application of this test method when estimating starting doses for acute toxicity testing but do find this test method to be a desirable addition to information about compound properties and related compounds.