2. How will the work contribute to further international harmonisation of hazard and risk assessment? Please provide details below or as an attachment.

Hazard identification and risk assessment of chemicals concerned with ocular irritation in human are very important in the development, control or regulation of chemicals.

For many years, the ocular irritation potential of chemicals mostly has been evaluated by the Draize test. As these test can involve pain and distress, it is desirable to find alternative methods that do not require the use of animals. Data obtained by *in vitro* tests can sometimes be obtained faster, and therefore serve an important role in the chemical risk assessment.

In light of animal welfare considerations, it is also important to harmonize available alternatives to animal testing.

or as attachment No.__

3. How will the proposed project address issues and /or endpoints which are of major human health or environmental concerns? If there are existing Test Guidelines or projects in the work plan of the Test Guidelines Programme covering the same endpoint, please refer to these and describe the added value and usability of the proposed new test method. Please provide details below or as an attachment.

Ocular irritation is a reaction caused by the direct contact of a chemical substance with the eye, inducing symptoms such as clouding of the cornea, inflammation of the iris, and redness/edema/secretion of the conjunctiva. It is important to assess ocular irritation, especially in products used on the face (such as cosmetics) or hair or household products, any of which can accidentally enter the eye.

The STE test uses corneal cells since the cornea is one of the main targets during accidental eye exposures, and damage to the cornea can result in visual impairment or loss. In addition, corneal effects are weighted heavily in the original *in vivo* ocular irritancy scoring systems (e.g., 80 out of a possible 110 points in the Draize eye test scoring system).

The endpoint evaluated in the STE test to measure the extent of damage to the SIRC cells following exposure to a chemical substance is cytotoxicity. Cytotoxicity is measured quantitatively based on relative viability of treated SIRC cells. Cell viability is measured by MTT assay method and a decrease in cell viability is correlated to corneal damage induced by ocular irritants.

The STE test is currently used in-house by Kao Corporation to assess the ocular irritation potential of industrial chemicals, cosmetics and personal care products.

or as attachment No.__

4. Will the project have general support from OECD member countries or is the outcome relevant for just one or a few member countries / stakeholders? Provide details of the countries and the rationale for this view below.

X Many countries ☐ A few countries ☐ Only for the submitting country

.The project is expected to be of interest and supported by most if not all OECD countries. an international peer review of STE test will be conducted by the end of 2011 under the coordination of NICEATM/ICCVAM and JaCVAM with accordance to the ICATM (International Cooperation on Alternative Test Methods) framework.

5. If the Test Guideline is not intended for general use, indicate if the Test Guideline would be intended for:

☐ Specific (limited) applications such as pesticide usage, or

☐ for specific classes of chemicals (e.g. surfactants) rather than for chemicals in general.

6. If the expected outcome of this proposal is a Test Guideline or a Guidance Document, provide information on the intended use, applicability and limitations of the test method.

The STE test using cultured corneal cells has the advantage of being a simple, quick procedure, and a low evaluation cost. Furthermore, poorly water-soluble chemicals like toluene, octanol, and hexanol could be evaluated in the STE test by using mineral oil as the vehicle (attachment No.1, ref.1). However, while a wide range of substances with various physicochemical characteristics can be tested in the STE test, substances that are insoluble either in saline, 5% DMSO, or mineral oil cannot be tested in the STE. In addition, colored test substances may be problematic as they could interfere with the optical density measured in MTT assay.

7. Provide supporting information on the validation status (i.e. relevance and reliability) of the method. Principles for validation of test methods for OECD Test Guidelines are described in Guidance Document 34.

Provide justification and rationale for the test, including data.

If there are no or limited data available to support the reliability and relevance of the proposed test, indicate if validation work is included in the project.

If there is no need for validation, provide a detailed justification.

Before starting the formal validation studies, inter-laboratory collaborative studies were performed in 2006-2009 (ref. 2, 3).

From the results of the abovementioned inter-laboratory collaborative study and validation studies conducted, it was demonstrated that the STE test was reproducible and applicable to detection of mild to severe ranges ocular irritants (attachment No.2, ref. 2, 3).

8. Describe if the test method includes components, equipment or other scientific procedures that are covered (or pending) by Intellectual Property Rights (IPR) (e.g., patents, patent applications, industrial designs and trademarks). Information should be provided on the overall availability of the IPR-protected components including whether they are commercially available or require a Material Transfer Agreement (MTA) or other licensing agreements. In addition, the possibility of providing a generic description of the IPR-covered component/test system should be disclosed and whether Performance Standards have been developed for the test method.

The proposed protocol includes no IPR-covered components/test system.

ADDITIONAL INFORMATION

In this section please provide further information to allow the Working Group of National Coordinators of the Test Guidelines Programme to assess the suitability of the project for the workplan of the Test Guidelines Programme

1. If the expected outcome of the project proposal is a Test Guideline and is based on existing, regional or international documents such as guidelines, protocols or guidance material, please provide that information here or as an attachment.

The expected outcome of the project proposal is a Test Guideline, but is not based on any existing, regional or international documents.

or as attachment No.__

2. If Animal Welfare considerations are addressed in the project proposal, provide details below or as an attachment. Explain if the project is aimed at refining, reducing and/or replacing the use of animals.

If the project is not specifically developed for animal welfare purposes, indicate if the animal welfare considerations have been a component of the project proposal.

Indicate if animal welfare considerations are irrelevant to the project, for example for physico-chemical properties.

In the Draize test, up to three animals for each test chemical are used and these animals are observed for a maximum 21 days post administration of the test substance to determine reversibility of effects. For the assessment of ocular irritation, one *in vitro* alternative test may not completely replace the Draize test. Therefore, a testing battery combining several *in vitro* assays, including cytotoxicity test, is proposed in order to estimate the entire irritancy range for chemicals. For these reasons, the STE test could be considered a component of this testing battery approach to establishing an ocular irritation animal alternative testing paradigm. Therefore, the STE test with other *in vitro* ocular irritation tests (e.g., BCOP, ICE) will contribute to a testing battery that may eventually replace the use of animals for estimation of ocular irritation potential of chemicals.

or as attachment No.__

3. Provide information on expected or possible resource savings in member countries as a result of this project.

Generally, cytotoxicity tests using cultured cells have the advantage of being simple, quick, and relatively inexpensive. The STE test was also well characterized in terms of employing cultured cells lines derived from cornea and possessing similar or shorter exposure times than many other cytotoxicity based methods (e.g., NRR, RBC).

4. If the expected outcome of the proposed project is a Guidance Document or Detailed Review Paper, will it be directly linked to the development of a particular Test Guideline or a series of Test Guidelines?

☐ Yes, it is the initial step in the development of a new or revision of existing Guidelines.

☐ Yes, additional guidance is needed for the most appropriate selection of the Guidelines on the subject.

☐ No, the guidance is on issues related to testing or the development of Test Guidelines in general.

There are 4 attachments added to this form.

Attachment No.1: Protocol of the STE test
Attachment No.2: Validation study report of the STE test (1st)
Attachment No.3: Validation study report of the STE test (2nd)
Attachment No.4: Published papers on the STE test (1998-2011)

ASSESSMENT OF PROJECT PROPOSAL

(To be completed by all member countries /stakeholders except the submitter)

Country / Organisation:	

Representative: (Preferably NC):	

Taking into account the project information, requested above, does this project meet the needs of the member countries for addition to the workplan of the Test Guidelines Programme

☐ Yes ☐ No ☐ Further information needed

If the response is "No" or "Further information needed", please provide justification:

Remarks as appropriate, including further information needs, if any:

別紙４

研究成果の刊行に関する一覧表

雑誌

発表者氏名	論文タイトル名	発表誌名	巻号	ページ	出版年
Hajime Kojima	Commentary to the Discussion on Topic 3, "In Vitro Test Approaches with Better Predictivity" at the 5th International Workshop on Genotoxicity Testing (IWGT)	Genes and Environment	Vol. 32 No. 2	40-42	2010
Hajime Kojima, Masahiro Takeyoshi, Takashi Sozu, Takumi Awogi, Kazunori Arima, Kenji Idehara, Yoshiaki Ikarashi, Yukiko Kanazawa, Eiji Maki, Takashi Omori, Atsuko Yuasa and Isao Yoshimura	Inter-laboratory validation of the modified murine local lymph node assay based on 5-bromo-2'-deoxyuridine incorporation	Journal of Applied Toxicology	Vol. 31	63-74	2010
小島 肇夫	動物実験の３Ｒにおける国内外の動向	ドージンニュース	No. 138	1-9	2011
廣川景子、亀山梨奈、中川真実子、井上智子、安部正道、稲葉弥寿子、山北高志、森敏恵、鈴木加余子、松永佳世子	香粧品パッチテスト2007年のまとめ	日本皮膚アレルギー・接触皮膚炎学会雑誌	Vol. 4 No. 2	89-98	2010
Akaza N, Akamatsu H, Sasaki Y, Takeoka S, Kishi M, Mizutani H, Sano A, Hirokawa K, Nakata S and Matsunaga K	Cutaneous Malassezia Microbiota in Atopic Dermatitis Patients Differ by Gender and Body Part.	*J Dermatol*	221	253-260	2010
Akaza N, Akamatsu H, Sasaki Y, Takeoka S, Kis	Cutaneous Malassezia microbiota of healthy subjects differ by sex, body par	*J Dermatol*	Vol. 37	786-792	2010

hi M, Mizutani H, Sano A, Hirokawa K, Nakata S, Matsunaga K	t and season.				
Furue M, Yamazaki S, Jimbow K, Tsuchida T, Amagai M, Tanaka T, Matsunaga K, Muto M, Morita E, Akiyama M, Soma Y, Terui T and Manabe M.	Prevalence of dermatological disorders in Japan: A nationwide, cross-sectional, seasonal, multicenter, hospital-based study	*J Dermatol*	Vol. 38	310-320	2010
Miura M, Isami M, Yagami A, Matsunga K.	Allergic contact cheilitis caused by ditrimethylol propane triethylhexanoate in a lipstick	Contact Dermatitis	Vol. 64 No. 5	301-302	2011
Hirota M, Motoyama A, Suzuki M, Yanagi M, Kitagaki M, Kouzuki H, Hagino S, Itagaki H, Sasa H, Kagatani S and AibaS.	Changes of cell-surface thiols and intracellular signaling in human monocytic cell line THP-1 treated with diphenylcyclopropenone.	J Toxicol Sci	Vol. 35 No. 6	871-879	2010
Hagino S, Okazaki Y, Kitagaki M and Itagaki H.	Further verification of an in vitro tier system for the identification of cosmetic ingredients that are not ocular irritants.	Altern Lab Anim	Vol. 38	139-152	2010
Ashikaga T, Sakaguchi H, Sono S, Kosaka N, Ishikawa M, Nukada Y, Miyazawa M, Ito Y, Nishiyama N and Itagak H.	A comparative evaluation of in vitro skin sensitisation tests: the human cell-line activation test (h-CLAT) versus the local lymph node assay (LLNA).	Altern Lab Anim	Vol. 38	275-284	2010
Kagatani S, Sasaki Y, Hirota M Mizuashi M, Suzuki M, Ohtani T, Itagaki H and Aiba S.	Oxidation of Cell Surface Thiol Grouops by Contact Sensitizers Triggers the Maturation of Dendritic Cells	The Journal of Investigative Dermatology	130	175-183	2010
Kano S, Todo H Sugie K, Fujimoto H, Nakada K, Tokudome Y, Hashimoto F and Sugibayashi K.	Utilization of Reconstructed Cultured Human Skin models as an Alternative Skin for Permeation Studies of Chemical Compounds	AATEX.	Vol. 15	61-70	2010

Todo H, Kimura E, Yasuno H, Tokudome Y, Hashimoto F, Ikarashi Y and Sugibayashi K.	Permeation pathway of macromolecules and nanospheres through skin.	Biol. Pharm. Bull.	Vol. 33	1394-99	2010
Onoue S, Takahashi H, Kawabata Y, Seto ., Hatanaka J, Timmermann B and Yamada S.	Formulation design and photochemical studies on nanocrystal solid dispersion of curcumin with improved oral bioavailability.	Journal of Pharmaceutical Sciences	Vol. 99	1871-81	2010
Kawabata Y, Yamamoto K, Debari K, Onoue S and Yamada S.	Novel crystalline solid dispersion of tranilast with high photostability and improved oral bioavailability.	European Journal of Pharmaceutical Sciences	Vol. 39	256-262	2010
Onoue S, Ochi M, Gandy G, Seto Y, Igarashi N, Yamauchi Y and Yamada S	High-throughput screening system for identifying phototoxic potential of drug candidates based on derivatives of reactive oxygen metabolites.	Pharmaceutical Research	Vol. 27	1610-19	2010
Seto Y, Ochi M, Onoue S and Yamada S	High-throughput screening strategy for photogenotoxic potential of pharmaceutical substances using fluorescent intercalating dye.	Journal of Pharmaceutical and Biomedical Analysis,	Vol. 52	781-786	2010
Seto Y, Ochi M, Igarashi N, Inoue R, Oishi A, Toida T, Yamada S and Onoue S.	*In Vitro* Photobiochemical Characterization of Sulfobutylether-β-cyclodextrin Formulation of Bufexamac.	Journal of Pharmaceutical and Biomedical Analysis,	Vol. 55	591-596	2010
尾上 誠良	最前線：薬剤性光線過敏症	ファルマシア	Vol. 47	295-300	2010
Yamamoto N, Hirano K, Kojima H, Sumitomo M, Yamashita H, Ayaki M, Taniguchi K, Tanikawa A and Horiguchi M.	Cultured human corneal epithelial stem/progenitor cells derived from the corneal limbus	In Vitro Cell. Dev. Biol. Anim.	Vol. 46 No. 9	774-780	2010

Genes and Environment, Vol. 32, No. 2 pp. 40–42 (2010)

Letter to the editor

Commentary to the Discussion on Topic 3, "*In Vitro* Test Approaches with Better Predictivity" at the 5th International Workshop on Genotoxicity Testing (IWGT)

Hajime Kojima[1]

Japanese Center for the Validation of Alternative Methods (JaCVAM), National Institute of Health Sciences, Tokyo, Japan

(Received March 17, 2010; Revised April 20, 2010; Accepted April 21, 2010)

The 5th International Workshop on Genotoxicity Testing (IWGT) was held on August 17–19, 2009, at Biozentrum of the University of Basel, Switzerland, prior to the 2009 International Conference on Environmental Mutagens (ICEM), Florence, Italy. In this workshop, approximately 200 participants from government, industry and academia, mainly from the USA, EU, Canada and Japan, discussed the following six topics: 1) suitable maximum concentrations for tests with mammalian cells (this group was divided into two subgroups: 1a) *in vitro* chromosome aberrations and micronuclei, and 1b) mammalian cell gene mutations); 2) photogenotoxicity testing requirements; 3) *in vitro* test approaches with better predictivity; 4) improvement of *in vivo* genotoxicity assessment, i.e., the link to standard toxicity testing; 5) use of historical control data in the interpretation of positive results; and 6) suitable follow-up risk assessment testing for *in vivo* positive results.

Dr. Toshio Kasamatsu (Kao Co.) and I were invited to the group on topic 3, "*In Vitro* Test Approaches with Better Predictivity". Over the course of seven presentations within the group, we discussed new test methods based on the following background. In sharp contrast with other groups, a new subgroup (led by Stefan Pfuhler, Procter & Gamble: P&G) was convened to develop consensus recommendations for choosing a better test systems to improve the predictivity of *in vitro* tests. This group reviewed current studies investigating whether certain cell types were more susceptible to give irrelevant positive results *in vitro*.

The background for the discussion is as follows:

1) The high rate of false positive results in the current battery of *in vitro* tests–as high as 80% when mammalian cell assays are combined (i.e., chromosome aberration assay and mouse lymphoma assay) (1);
2) Seventh Amendment to EU Cosmetics Directive: Marketing and testing ban on ingredients tested *in vivo*, which came into force March 2009 (2);
3) REACH (Registration, Evaluation, Authorization and Restriction of Chemical) issue. The REACH Regulation gives greater responsibility to industry to manage the risks from chemicals and to provide safety information on the substances (3).

Particularly, the European testing ban resulting from the Seventh Amendment to the EU Cosmetics Directive may result in valuable compounds being unnecessarily discarded from European cosmetic markets. In the field of mutagenicity/genotoxicity, validated alternative methods are available, and *in vivo* studies were prohibited in EU territory after March 11, 2009 (2). In spite of the high rate of false positive results in the current battery of *in vitro* tests, the Scientific Committee on Consumer Safety (SCCS) (4), a European regulatory agency, recommends a battery of three *in vitro* assays: a bacterial reverse mutation test; an *in vitro* mammalian cell gene mutation test; and either an *in vitro* micronucleus test or *in vitro* mammalian chromosome aberration test. False positive results may be due to experimental conditions that have no relevance to *in vivo* situations. It is hard to confirm or deny the genotoxicity potential of candidate cosmetic ingredients based on results of the current *in vitro* tests. Therefore, we hope to develop promising new approaches for *in vitro* testing that can be used in place of *in vivo* studies to reduce the high incidence of false positive results of existing *in vitro* testing.

First, the group addressed the following questions:

Choice of cell line:

→ Are there restrictions on the choice of cells used for testing?
→ Should there be recommendations about the choice of cell line? Can this recommendation be made now, or are more data needed?

[1]Correspondence to: Hajime Kojima, Japanese Center for the Validation of Alternative Methods (JaCVAM), National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan. Tel: +81-3-3700-9874, Fax: +81-3-3700-1145, E-mail: h-kojima@nihs.go.jp

New approaches:

→ How do we rate the status of the genotoxicity assays performed on 3-dimensional human skin equivalents (3-D skin)?

→ What is the applicability domain of that assay?

→ What kind of data would the team like to see before considering this as a valid assay?

Presentations

1) Which cell lines should we be using?

✓ Paul Fowler (Covance) presented "Reduction of false positives in *in vitro* genetic toxicology testing: Importance of cell selection"; the "false positives" project, which compares rates of false positives with V79, CHO, CHL, TK6, HepG2, L5178Y, and HuLy obtained from COLIPA; European trade association for cosmetic, toiletry and perfumery industry;

✓ Mick Fellows (AstraZeneca) presented "Data generated in conjunction with the *in vitro* micronucleus test (MNvit) OECD guideline finalisation";

✓ Azeddine Elhajouji (Novartis) presented "Comparison of V79, L5178Y, TK6 and primary human lymphocytes for micronucleus induction"; and

✓ Ludowig Le Hegarat (Agence Francaise De Securite Sanitaire Des Aliments) presented "The HepaRG cell line, a predictive model in genetic toxicology?"

2) Promising new approaches

✓ Rodger Curren (Institute for *In vitro* Science: IIVS) presented "A novel micronucleus assay conducted in reconstructed human skin"; micronucleus data generated in 3-D skins (COLIPA 3-D skin project);

✓ Gladys Ouédraogo (L'Oreal) presented "The comet assay on 3-D skins (COLIPA 3-D skin project)";

✓ Hajime Kojima presented "Proposal of protocol for comet assay using a three-dimensional epidermal model".

From these presentations, we reached a consensus, which were reflected in the following statements in the final IWGT comments:

1. Data were presented indicating that p53-compromised rodent cell lines over-estimate genotoxic potential in the micronucleus test. Therefore, IWGT suggests *in vitro* micronucleus or chromosome aberration assay using p53-competent cells.

The data presented at the IWGT from the OECD MNvit test showed that all cell types correctly identify clastogens and aneugens. However, the data of the compounds that gave false results in the MNvit test suggested that there was a great diversity of the responses of the various cell types. The group agreed, based on the similarity of MNvit and chromosome aberration assays, that the diversity would also be observed in the chromosome aberration assay. HepaRG is a promising model, in that the cells appear to have better phase I and II metabolizing potential than the other cell lines. However, further evaluation is required in order to confirm the value of this model for genotoxicity testing.

2. It has been demonstrated that cell line stability and source can affect the outcome of genotoxicity assays. Therefore, IWGT recommends adhering to good cell culture practices, characterizing all new cells, checking regularly for drift, and working from low-passage stocks. It would be useful to compile a common genotoxicity cell bank with fully characterized stocks of all cells.
3. Genotoxicity testing in 3-D skins (micronucleus test and comet assay) is a promising new *in vitro* test for chemicals applied to the skin.

The advantage of the model is that it resembles the properties of human skin (barrier function, metabolism), and that the route of exposure is relevant for dermally applied chemicals (e.g., cosmetics). The data presented show that the micronucleus assay by 3-D skin is further advanced; furthermore, inter-and intra-lab reproducibility has been demonstrated. IWGT agreed that the comet assay should be further evaluated. The comet assay is seen as a valuable addition, as it is not dependent on cell proliferation and covers a wider spectrum of DNA damage. The metabolic capacity needs to be further evaluated (this work is ongoing). It would be valuable to capture the kinetics of penetration and toxicity in order to establish the ideal sampling time(s) for the comet assay. Recommendations on the use of appropriate vehicles should be established. It was agreed that 3-D skin, once validated, will be useful for following up on positive results from standard *in vitro* assays for dermally applied chemicals. The applicability domain will be established once the validation is completed.

Considering the EU situation and the international expansion of animal welfare laws, we must avoid progressively adopting *in vivo* genotoxicity testing. Therefore, I think that the consensus in this group was made "according to script", at least with regard to European colleagues. Against European drastic transformation, my concern on the safety evaluation is that the trend may push out *in vivo* testing in the cosmetic field. We must send out a warning about the current trend to avoid all *in vivo* tests and emphasize importance of re-evaluating cell lines and validating studies on new approaches to testing *in vitro*. I think further discussion is needed to established well-balanced *in vitro* and *in vivo* tests for evaluation of genotoxic risk of chemicals to human.

References

1 Kirkland D, Kasper P, Muller L, Corvi R, Speit G. Recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests: A follow-up to an ECVAM workshop. Mutat Res. 2008; 653: 99–108.

2 Commission Staff Working Documents, Time tables for the phasing-out of animal testing in the framework of the 7th Amendment to the Cosmetics Directive (Council Directive 76/768/EEC), EN, SEC 2004; 1210: 1–8.

3 JRC. 2008. http://www.vet.uu.nl/nca/userfiles/other/REACH

4 Scientific Committee on Consumer Safety (SCCS)/1294/10, Memorandum on "Alternative test methods in human health safety assessment of cosmetic ingredients in the European Union". 2009.

Research Article

Journal of Applied Toxicology

Received: 10 April 2010, Revised: 27 May 2010, Accepted: 28 May 2010 Published online in Wiley Online Library: 30 July 2010

(wileyonlinelibrary.com) DOI 10.1002/jat.1567

Inter-laboratory validation of the modified murine local lymph node assay based on 5-bromo-2′-deoxyuridine incorporation

Hajime Kojima,[a]* Masahiro Takeyoshi,[b] Takashi Sozu,[c] Takumi Awogi,[d] Kazunori Arima,[e] Kenji Idehara,[f] Yoshiaki Ikarashi,[a] Yukiko Kanazawa,[g] Eiji Maki,[h] Takashi Omori,[i] Atsuko Yuasa[j] and Isao Yoshimura[k]

ABSTRACT: The murine local lymph node assay (LLNA) is a well-established alternative to the guinea pig maximization test (GPMT) or Buehler test (BT) for the assessment of the skin sensitizing ability of a drug, cosmetic material, pesticide or industrial chemical. Instead of radioisotope using in this method, Takeyoshi M. *et al.* (2001) has developed a modified LLNA based on the 5-bromo-2′-deoxyuridine (BrdU) incorporation (LLNA:BrdU-ELISA). The LLNA:BrdU-ELISA is practically identical to the LLNA methodology excluding the use of BrdU, for which a single intraperitoneal injection of BrdU is made on day 4, and colorimetric detection of cell turnover. We conducted the validation study to evaluate the reliability and relevance of LLNA:BrdU-ELISA.

The experiment involved 7 laboratories, wherein 10 chemicals were examined under blinded conditions. In this study, 3 chemicals were examined in all laboratories and the remaining 7 were examined in 3 laboratories. The data were expressed as the BrdU incorporation using an ELISA method for each group, and the stimulation index (SI) for each chemical-treated group was determined as the increase in the BrdU incorporation relative to the concurrent vehicle control group. An SI of 2 was set as the cut-off value for exhibiting skin sensitization activity.

The results obtained in the experiments conducted for all 10 chemicals were sufficiently consistent with small variations in their SI values. The sensitivity, specificity, and accuracy of LLNA:BrdU-ELISA against those of GPMT/BT were 7/7 (100%), 3/3 (100%), and 10/10 (100%), respectively. Copyright © 2010 John Wiley & Sons, Ltd.

Keywords: 5-bromo-2′-deoxyuridine (BrdU); inter-laboratory validation study; local lymph node assay (LLNA); skin sensitization; stimulation Index (SI)

INTRODUCTION

Skin sensitization (i.e. allergic contact dermatitis, ACD) is an immunologically mediated cutaneous reaction to a drug, cosmetic material, pesticide or industrial chemical. It is known that the detection and evaluation of the immune-based adverse effects that are collectively referred to as hypersensitivity reactions are very difficult tasks, particularly during the drug approval process, because of the lack of adequate non-clinical models and the low incidence rate of reactions (Hastings, 2001). However, there are several adequate and predictive methods for modeling ACD. For several decades, tests involving guinea pigs, such as the guinea pig maximization test (GPMT) or the Buehler test (BT), have been used for assessing the skin sensitization potential of chemicals (OECD, 1992). In addition, a mouse model for assessing the relative sensitization potential is a well-established alternative method for determining whether a chemical causes ACD. Although GPMT and BT can be viewed as phenomenological methods in which the clinical signs are modeled, local lymph node assay (LLNA) and the mouse ear swelling test are based on a mechanistic understanding of immune-based contact dermatitis (Gad *et al.*, 1986; Hastings, 2001). In addition, these methods also offer important animal welfare benefits. In these assays, the use of LLNA has been successfully validated by several studies (Basketter and Scholes, 1992; Basketter *et al.*, 1996, 2002; Gerberick *et al.*, 2000; Haneke *et al.*, 2001). Recently, it has been formally adopted by the Organization for Economic Co-operation and Development (OECD), according to the guidelines for testing

*Correspondence to: *H. Kojima, Japanese Center for the Validation of Alternative Methods, National Institute of Health Sciences, 1-18-1 Kamiyoga, Setagaya-ku, Tokyo 158-8501, Japan.*
E-mail: h-kojima@nihs.go.jp

[a]*National Institute of Health Sciences, Tokyo, Japan*

[b]*Chemicals Evaluation and Research Institute, Kitakatsushika, Saitama, Japan*

[c]*Osaka University, Osaka, Japan*

[d]*Otsuka Pharmaceutical Co. Ltd, Naruto, Tokushima, Japan*

[e]*Taisho Pharmaceutical Co. Ltd, Saitama, Japan*

[f]*Daicel Chemical Industries Ltd, Himeji, Hyogo, Japan*

[g]*Food and Drug Safety Center, Hadano, Kanagawa, Japan*

[h]*Biosafety Research Center, Foods, Drugs and Pesticides, Iwata, Shizuoka, Japan*

[i]*Kyoto University School of Public Health, Kyoto, Japan*

[j]*Fuji Film Co. Ltd, Ashigara, Kanagawa, Japan*

[k]*Tokyo University of Science, Tokyo, Japan*

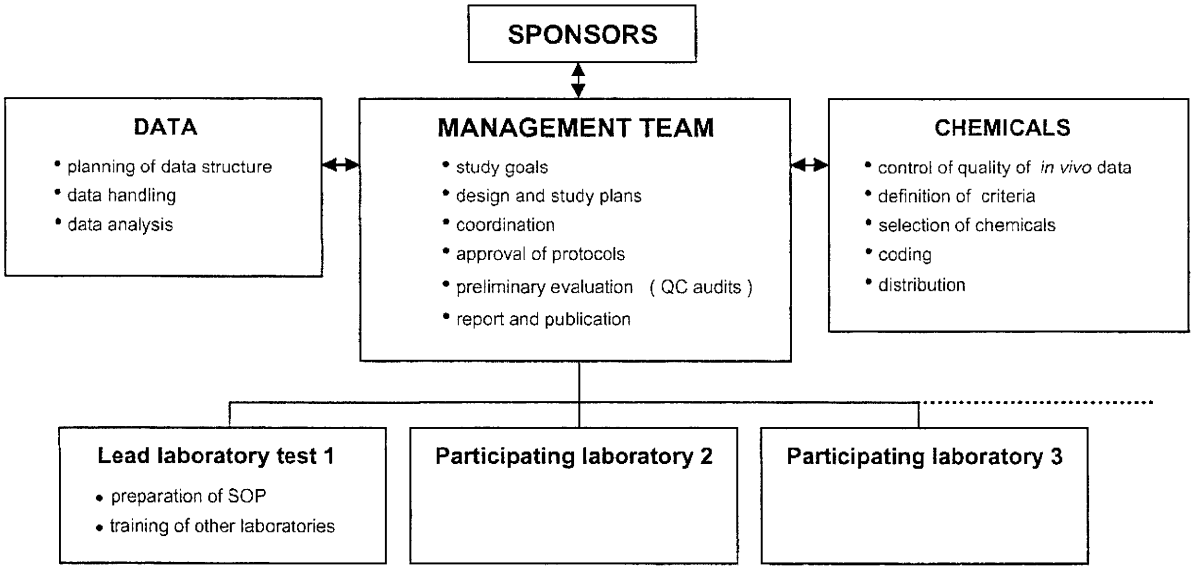

Figure 1. The organization of validation management structure.

chemicals 406 and 429 (OECD, 1992, 2002), and it is accepted by the EU and USA as a suitable method for classifying the skin sensitizing ability of chemicals (Basketter *et al.*, 2005; Dean *et al.*, 2001; Sailstad *et al.*, 2001). The LLNA is specifically designed to identify contact allergens. The assay is not intended to facilitate the detection of low molecular weight chemicals associated with systemic sensitization or drug allergies (Kimber, 2001). However, an investigation, which was designed to explore the ability of LLNA to identify pharmaceutical process intermediates known to cause contact allergy in humans, provided evidence that the assay is a useful method for hazard identification (Durand *et al.*, 2003). Furthermore, the use of the method, along with the use of GPMT and BT, is recommended for the determination of the skin sensitization potential of new drugs (FDA, 2002).

The original LLNA uses [3H]-methyl thymidine to measure lymphocyte proliferation. The use of this method is hindered, particularly in Japan, because such a radioisotope (RI)-based method requires special facilities and handling procedures. Several authors have been conducting investigations for the development of an alternative non-RI method for performing LLNA (Dearman *et al.*, 1999; Ehling *et al.*, 2005a, b; Hatao *et al.*, 1995; Idehara *et al.*, 2008, Lee *et al.*, 2002; Omori *et al.*, 2008; Takeyoshi *et al.*, 2001, 2005, 2006; Yamashita *et al.*, 2005).

One method, the LLNA:BrdU-ELISA proposed by Takeyoshi *et al.*, is a modification of the original LLNA that involves assessing lymphocyte proliferation using ELISA (enzyme-linked immunosorbent assay) to measure 5-bromo-2′-deoxyuridine (BrdU) incorporation instead of measuring the radioactivity produced by [3H]-methyl thymidine incorporation. This modification is also one of the most promising non-radioisotopic LLNA modifications that has been recently peer-reviewed by the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) Peer Review Panel Meeting for New Versions and Applications of the Murine Local Lymph Node Assay (ICCVAM 2009).

Although the methodology of the LLNA:BrdU-ELISA is essentially the same as the original LLNA, the data available were insufficient for validating the assay method for inter-laboratory reproducibility. Therefore, we conducted the inter-laboratory validation studies for LLNA:BrdU-ELISA. Prior to the studies, a preliminary study was conducted using only the positive control chemical, namely, 50% hexyl cinnamic aldehyde (HCA). After that, the first study testing 12 chemicals in nine laboratories yielded a large variation among tests. It was considered that extreme variation in the stimulation index (SI) was produced by the very small BrdU labeling index of the negative control. Then we determined that protocol revisions were necessary. We describe the phase II study in which the revised protocol was used. The objectives of the study were (1) to evaluate the extent of inter-laboratory variation with regard to LLNA:BrdU-ELISA and (2) to ascertain whether the results of LLNA:BrdU-ELISA are comparable with those of LLNA and GPMT/BT.

METHODS

Organization

This study was organized by researchers belonging to the validation management team (VMT) for the validation of the assay. The structure of the VMT is shown in Fig. 1. The VMT comprised representatives from each experimental laboratory, toxicologists as the chemical selectors and distributors of the chemicals and materials, biostatisticians and the study manager. All the experiments were performed by the toxicologists of the experimental laboratories. In this study, nine experimental laboratories with sufficient experience in the use of the LLNA and/or its modifications participated. Research teams of all the experimental laboratories obtained ethical approval for each standard operational procedure conducted in their laboratories. After the preliminary and pre-validation studies by the VMT, seven laboratories, excluding Chemicals Evaluation and Research Institute, Japan (CERI) and Meiji Seika Kaisha Ltd, participated in the final validation study.

Journal of AppliedToxicology

Technology Transfer

A one-day technology-transfer seminar was held by the LLNA:BrdU-ELISA developer, which was attended by at least one researcher (or technician) from each experimental laboratory. Participants learned the method for conducting the assay according to the standard protocol.

Test Phases and Schedule

Prior to the study, a preliminary study was conducted by researchers from all the experimental laboratories, whic used only the positive control chemical, 50% HCA. Two phases of the validation study were conducted. In phase I (pre-validation study), we examined the reliability and the reproducibility of the test protocol. Inter- and intra-laboratory variability and transferability were assessed using suitable statistical analysis from testing 12 chemicals in nine laboratories from August to December 2006. However the variability of the SI values for the positive control was too large, the VMT determined that protocol revisions were necessary without analyzing the data for the 12 chemicals. This report details the phase II study, in which 10 new chemicals were tested by seven laboratories from September to December 2007 using the revised protocol.

Chemical Selection and Allocation

The chemical selectors chose 20 candidate chemicals that were previously used in LLNA and whose test results had been documented (Basketter and Scholes, 1992; Basketter *et al.*, 1998, 1999a, b, 2000; Gerberick *et al.*, 2004; Haneke *et al.*, 2001; Kimber *et al.*, 1998; Loveless *et al.*, 1996). On the basis of these literature data and solubility of the chemicals, the chemical selectors selected vehicles and prepared three fixed doses (low, medium and high) for each chemical. Subsequently, the chemicals were transported from the chemical and material distributors to the experimental laboratories.

In this study, 10 of the 20 candidate chemicals were selected and classified as strong, mild or weak sensitizers or non-sensitizers on the basis of original LLNA data. In order to reduce the number of animals used, pairs comprising groups treated with three chemicals and the corresponding vehicle control group were employed. In other words, in each laboratory, three chemicals were simultaneously tested with one negative control and one positive control for every experiment. Of the 10 chemicals, three were dispatched to all seven participating experimental laboratories, and the remaining seven were randomly allocated to the laboratories by a biostatistician and dispatched to three experimental laboratories.

So that the research teams could not predict the severity of the effects for each chemical, all the chemicals were coded into alphabetic characters for blinded distribution, and indicated as low, medium and high in terms of the dose. Each research team prepared solutions using the distributed solvents before experiments. However, prior to the study, the researchers and toxicologists of the respective laboratories were informed of the identity of the 20 candidate chemicals and the corresponding control vehicles. This was done in order to ensure the safety of the chemists performing the experiments (e.g. with regard to proper disposal of the chemicals) and to prevent any anxiety that they would experience while handling unknown chemicals.

Standard Protocol for LLNA:BrdU-ELISA

The standard protocol for the assay was prepared prior to the preliminary test and refined before the commencement of the study.

Animals

Young adult female mice (nulliparous and non-pregnant) of the CBA/JN strain were used at age 8–12 weeks. Healthy animals in good general condition on arrival were quarantined for more than 5 days. During the quarantine and the acclimation period, clinical signs, body weights and excrement of the animals were monitored. Animals confirmed to be in good health with favorable body weight gains by a person in charge of animal management during the quarantine and acclimation period were allocated to groups by a stratified randomization or other appropriate methods before the start of the study. Animals were identified by marking the tail with colored marker, ear tags or other appropriate methods.

The animals were housed in an animal room maintained at a temperature of 22 ± 3°C and a relative humidity of 30–70%. The rooms were artificially lighted for 12 h daily, and the animals were given free access to conventional laboratory diet and drinking water.

Animal experiment

A minimum of four successfully treated animals was used per dose group, with a minimum of three consecutive doses of the chemical, and one group each for the negative (vehicle) control and positive control. A 25 µl dose of test solution was applied to the dorsum of both ears of the mice for three consecutive days using a microvolume pipette. A single intraperitoneal injection of 0.5 ml of BrdU solution (5 mg/mouse/injection) was given to the mice 48 h after the final application. BrdU solution was prepared before administration and stored in a freezer below –20°C until use.

Clinical signs were observed at least once per day. Body weights were measured on the day of the first application and on the day lymph nodes were collected. Approximately 24 h after BrdU injection, the auricular lymph nodes were removed. The lymph nodes were carefully dissected and trimmed of fascia and fat, weighed and stored individually in a 1.5 ml centrifuge tube at –20°C until BrdU-ELISA measurement.

Preparation of Lymph Nodes (LN) Cell Suspension

From the discussion of the results on the phase I, we determined that the mean absorbance of the negative (vehicle) control group should be within 0.1–0.2. Because the absorbance depends on the combination of assay apparatus and the target volume of cell suspension, every laboratory had to decide their own optimal target volume of LN cell suspension in advance so that the absorbance value of the negative control group would be within 0.1–0.2. The volume would be expected to be approximately 15 ml. The volume of LN cell suspension of the all test animals was adjusted to the optimized volume.

A small amount (*ca* 0.3 ml) of physiological saline was added to the centrifuge tube that contained the collected LN, and the LN was crushed with a disposal plastic pestle to make the cell

suspension. The cell suspension was passed through a #70 nylon mesh and adjusted to the target volume in a 50 ml tube.

Assay Flow (BrdU-ELISA)

The incorporation of BrdU into LN cells was determined using a commercial cell proliferation assay kit (Roche Diagnostics GmbH, Roche Applied Science, 68298 Mannheim, Germany; catalog no. 11 647 229 001) after the lymph node was crushed and suspended in physiological saline. The absorbance was defined as the BrdU labeling index.

The cell suspension (100 μl) was added to the wells of a flat-bottom microplate (three wells per sample) after being homogenized by mixing thoroughly with a Vortex mixer. Simultaneously, three blank wells were prepared by adding 100 μl of physiological saline. After adding all samples and blank solution to wells, the plate was centrifuged at 300 *g* for 10 min.

Three-quarters of the volume of the supernatants was removed. In this step, great care was taken not to aspirate the LN cells. The assay plate was dried completely in a hot-air oven.

Two hundred microliters of Fix-Denat solution was added into each well and allowed to stand for 30 min at room temperature. Then the Fix-Denat solution was completely removed.

One hundred microliters of anti-BrdU antibody working solution was added and allowed to stand for 1 h. Then the anti-BrdU antibody working solution was completely removed.

Two hundred microliters of wash solution was added to each well, and the well was washed by pippetting 10 times. Then the wash solution was completely discarded. The wash step was repeated twice (three times in total).

One hundred microliters of tetramethylbenzidine substrate solution was added and allowed to stand for 15 min at room temperature in a dark place. The absorbance at 370 nm was measured with a reference wavelength of 492 nm. When using stop solution (1 M sulfuric acid, 25 μl/well), the absorbance at 450 nm was measured with a reference wavelength of 690 nm.

The BrdU labeling index and SI are defined as follows: without stop solution, BrdU labeling index = (ABS370-ABSblank370) – (ABS490-ABSblank490); with stop solution, BrdU labeling index = (ABS450-ABSblank450) – (ABS650-ABSblank650). The BrdU labeling index was determined for each test animal. SI is the ratio of the mean BrdU labeling index for the treated group to the mean for the concurrent vehicle control group

The positive/negative decision was made based on the criteria of SI ≥ 2, rather than SI ≥ 3, because the dynamic range of the LLNA:BrdU-ELISA is narrower than scintillation counting of incorporated [3H]-methyl thymidine and it produces lower SI values compared with the original LLNA.

Acceptance criteria for each experiment

Under the optimized assay conditions described, the mean absorbance range of the negative control is 0.1–0.2. Where the absorbance was higher than 0.2, the stock solutions of the assay were diluted and the absorbance was re-measured. Furthermore, the SI for the positive control group (50% HCA) was equal to or greater than 2. If not, data derived from the experiment were not employed for evaluation.

Adherence to Good Laboratory Practices

The studies were not conducted under full compliance with Good Laboratory Practices (GLP). However, all the laboratories were equipped to perform, and competent with, GLP. In addition, all the laboratories that participated in the inter-laboratory validation studies used the same experimental protocol and took part in a one-day seminar in which the protocol and execution of the test method were explained. Also, the same commercial kit, test materials and the same dose of each coded substance were used in all the laboratories.

Quality Control Check

A formatted file for the entry of the experimental data and information was prepared using Microsoft Excel. The file was distributed to the experimental laboratories prior to the experiment. After all testing was completed, all records and documents were checked by the chief and biostatistician of the VMT. If there were missing or strange points, the laboratory was asked for all documents to be submitted to the VMT.

Database

A biostatistician created a database containing the body weight on the start and final days, LN weight and BrdU incorporation data obtained from each mouse in all of the experimental laboratories. For comparison, data from studies on the original LLNA were collected and included in the database.

Statistical Methods

In order to demonstrate the variability within the SI values, the confidence interval of the SI values was calculated (Omori and Sozu, 2007). A variance component, τ^2, estimated by a random effect model for the log-transformed SI, was used as a measure of the inter-laboratory variations; this is similar to the meta analysis technique used in clinical studies (Normand, 1999). Using the above-mentioned random effect model, we estimated the weighted average as an overall estimate of the SI value for each chemical dose. The EC2 was defined as the estimated concentration that yielded an SI of 2 from the dose–response curve. If the response for an absorbance was not clearly dose-related, but the Si was greater than or equal to 2 for two doses, then it was considered to be positive. We defined this rule as the positive criterion. The EC2 of the weighted average was estimated and classified into the appropriate chemical category (Gerberick *et al.*, 2004). Finally, the sensitivity, specificity, accuracy, positive predictivity and negative predictivity were calculated as measures of relevance on the basis of the weighted averages in order to assess the concordance of the LLNA:BrdU-ELISA results with the LLNA or GPMT/BT results (OECD, 2005).

RESULTS

Chemical Selection

Table 1 shows the selected chemicals, the results of LLNA and GPMT/BT as references, and the LLNA:BrdU-ELISA results. The GPMT/BT results for chemical F (glutaraldehyde) are not listed in Table 1 because the data were not available at the time the list was prepared.

Body Weights

Table 2 summarizes the body weight statistics observed on days 1 and 6 in each laboratory. No substantial inter-laboratory varia-

Journal of Applied Toxicology

Table 1. Selected chemicals with their corresponding vehicles, the referenced results of LLNA and GPMT-BT, and the allocation of chemicals for the LLNA:BrdU-ELISA experiments

Code	Chemical	CASRN[a]	Vehicle[b]	LLNA	GPMT/BT[c]	Dose	Laboratory						
							1	2	3	4	5	6	7
A	Nickel sulfate [Nickel(II) sulfate hexahydrate]	10101-97-0	DMSO	+	+	1,3,10			○	○			○
B	Isopropanol (2-propanol)	67-63-0	AOO	−	−	10,25,50	○	○	○	○	○	○	○
C	Eugenol	97-53-0	AOO	+	+	10,25,50	○					○	○
D	trans-Cinnamic aldehyde	104-55-2	AOO	+	+	1,3,10							
E	2,4-Dinitrochlorobenzene (1-chloro-2,4-dinitrobenzene)	97-00-7	AOO	+	+	0.1,0.3,1.0	○	○	○	○	○	○	○
F	Glutaraldehyde solution (ab.25%)	111-30-8	ACE	+		0.1,0.3,1.0	○				○	○	
G	Methyl salicylate	119-36-8	AOO	−	−	10, 25,50	○	○	○				
H	Hexylcinnamic aldehyde (hexylcinnamal, α-Hexylcinnamaldehyde)	101-86-0	AOO	+	+	10, 25, 50	○	○	○	○	○	○	○
I	Lactic acid	598-82-3	DMSO	−	−	10, 25, 50		○	○				○
J	Formaldehyde solution (36–38%)	50-00-0	ACE	+	+	1,3,10	○				○		○

[a]The Chemical Abstract Services Registry Number.
[b]ACE, acetone; AOO, acetone–olive oil; DMSO, dimethylsulfoxide.
[c]Judgment based on the guinea pig maximization test or the Buehler test.

tions were observed with regard to the body weights. The lack of four data points at one laboratory (no.6) was due to experimental problems. There was no evidence that systemic toxicity occurred at any of the doses tested by reference to other observation records.

Acceptance Criteria

In the standard protocol, the mean absorbance in the negative control was set at 0.1–0.2 as acceptance criteria. However, approximately 60% of all data in this study were within this range and the other data were outside the range (data not shown). All the mean absorbances before the stock solutions were diluted were 0.05–0.35. Therefore, we discussed the specified range and changed the criteria. We accepted all data obtained before dilution for the following analysis.

Figure 2 shows the SI values obtained for all the positive control groups with 95% confidence intervals. Most experiments in these studies were acceptable because the SI values were greater than 2, excluding one laboratory (no. 2, test 1). Because all data at Laboratory no. 2, test 1, were judged to be insufficient for the positive control criteria, they were not used for the following analysis.

BrdU Labeling Index and SI Values

The BrdU labeling index and SI values recorded for each of the test chemicals in the experimental laboratories are summarized in Table 3, and the dose–response relationships for the SI values are indicated in Fig. 3. Dose–response relationships for the SI values of chemicals A (nickel sulfate), C (eugenol), D (*trans*-cinnamic aldehyde), E (2,4-dinitrochlorobenzene), F (glutaraldehyde) and H (HCA) were clearly evident in each laboratory, and the SI values at all their high doses were greater than 2. The dose–response relationships for chemicals B (isopropanol) and G (methyl salicylate) were unclear, and the laboratories that assessed these chemicals reported them to be negative. For chemical I (lactic acid), the dose–response relationships of the SI values were clearly evident at two laboratories (3 and 7), and the SI value for the high dose at one laboratory (7) was greater than 2. The dose–response at the other laboratory (4) was unclear. Further, an inconsistency was observed in the BrdU labeling index in the vehicle control group for one chemical. In the case of chemical J (formaldehyde), the dose–response relationship of the SI values yielded a down-regulation at the high dose in one laboratory (1), a v-shaped curve in one laboratory (5) and were unclear in one laboratory (6); therefore, it may be considered that the observed dose–response relationships based on the weighted average values for chemical J were inappropriate.

Intra-laboratory Variability

Although limited, the results obtained from the positive control groups allowed us to evaluate the intra-laboratory variability of the assay. Figure 2 shows the variability of the SI values obtained for the positive control groups from each laboratory. The SI values of the positive control substance were consistent, excluding the first trials of laboratories 2 and 5.

Inter-laboratory Variability

The data shown in Fig. 3 were used to measure the inter-laboratory variability of the SI values at all the chemical doses.

Table 2. Summary statistics of the body weight (g) at days 1 and 6

Laboratory	Day 1					Day 6				
no.	*n*	Mean	SD	Min	Max	*n*	Mean	SD	Min	Max
1	108	22.2	1.38	18.4	25.8	108	22.6	1.48	18.7	26.4
2	108	22.6	1.32	20.1	21.7	108	23.8	1.52	20.6	28.0
3	108	22.1	1.38	19.3	26.2	108	23.1	1.48	20.0	27.0
4	108	21.8	1.44	17.6	25.9	108	22.4	1.57	18.1	26.1
5	108	22.6	1.25	19.6	25.2	108	22.8	1.36	19.7	26.0
6	104	22.0	1.30	21.0	25.3	104	22.0	1.27	19.0	24.8
7	108	22.1	1.55	18.9	27.8	108	22.9	1.42	19.7	26.3

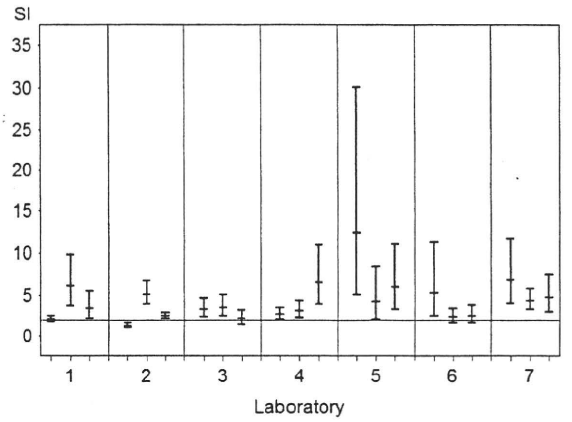

Figure 2. SI values with 95% confidence intervals obtained for the positive control groups at the all laboratories in phase II validation study.

Table 4 shows the weighted average of the SI values with 95% confidence intervals and a summary index of the inter-laboratory variability, i.e. τ^2. In the study, no large inter-laboratory variation was observed in any of the laboratories.

Table 5 shows the results of the judgments based on the positive criteria obtained for all the chemicals in all the laboratories. Two chemicals, I (lactic acid) and J (formaldehyde), showed inconsistent results among the laboratories. The SI values for the high doses were approximately 2 among all three laboratories; thus, the variation was small.

EC2 and Measures of Relevance

To avoid the problem of multiple counts of the same chemicals from different laboratories, the calculations of EC2 and sensitivity, specificity, accuracy, positive predictivity and negative predictivity of LLNA:BrdU-ELISA were based on the weighted averages of the SI values.

Table 6 show the EC2 results and its classification for LLNA:BrdU-ELISA based on the weighted averages for both the studies and the reported EC2 and its classification based on the reported values for the original LLNA. The sensitivity, specificity, accuracy, positive predictivity and negative predictivity of LLNA:BrdU-ELISA with regard to the chemicals in the study, as against those of GPMT/BT and LLNA, are shown in Table 7. To enable comparison of the measurements of LLNA against those of GPMT/BT when the same chemicals were selected, these values calculated only on the basis of the referenced data are shown in Table 7. The sensitivity, specificity, accuracy, positive predictivity and negative predictivity values of LLNA:BrdU-ELISA against those of GPMT/BT were similar to those of LLNA against those of GPMT/BT.

DISCUSSION

Through the preliminary and phase I studies, we developed and refined the protocol. For acceptable performance of the assay, the acceptance criteria as shown in the protocol must be used. That is, the appropriate absorbance for the negative and positive control must be obtained to achieve an adequate SI value of the positive control. To achieve this objective, the LN cell suspension should be prepared such that the mean absorbance on the negative and positive control is within the range defined by the protocol developers. However, even if the optimized condition was determined by a preliminary study in each laboratory, it was technically difficult to achieve the required absorbance range for the negative control in phase I of the study. As the result, absorbances on the negative and positive controls were in the ranges 0.002–0.398 and 0.003–1.115, respectively, in the phase I study (data not shown). Since the width of this range is so large, the range of average SI values on the positive control was also large, such as 1.7–30.2 at all laboratories.

To modify the acceptable range of the absorbance on the negative control, we discussed the use of data from the diluted stock of the suspension after the phase I study. It was suspected that the error was increased by dilution of the stock. We ultimately decided that the stock of LN cell suspension should not be used. The description of the dilution was eliminated.

Furthermore, we determined that every laboratory should decide their own optimal target volume of LN cell suspension in advance such that the absorbance value of the negative control group would be expected to be within 0.1–0.2.

Using the revised protocol that included these decisions, the phase II study was conducted with the participation of seven independent experimental laboratories. Unfortunately, absorbance data for the negative control did not meet the acceptance range of 0.1–0.2; the absorbance range was 0.05–0.35.

Because the results for the positive control were stable and the range of SI values was between 2.1 and 12.5 at seven laboratories, we accepted all the data for analysis in this study. We believe the results presented here are appropriate to understand the features of the assay.

The fact that the inter-laboratory variations were small for most of the chemicals is considered to be a significant finding of this study. In particular, chemical B (HCA) and chemical E (2,4-

Table 3. Mean absorbance and SI value

Code	Chemical name	Lab. no.	Dose	*n*	Mean absorbance for chemical	Mean absorbance for vehicle	SI
Positive	HCA	1	50	4	0.432	0.209	2.07
			50	4	0.337	0.055	6.11
			50	4	0.282	0.082	3.43
		2	fail				
			50	4	0.677	0.131	5.15
			50	4	0.438	0.174	2.52
		3	50	4	0.804	0.241	3.34
			50	4	0.720	0.203	3.54
			50	4	0.689	0.316	2.18
		4	50	4	0.756	0.281	2.69
			50	4	0.710	0.224	3.17
			50	4	1.012	0.154	6.58
		5	50	4	1.569	0.126	12.46
			50	4	0.683	0.161	4.24
			50	4	0.678	0.112	6.07
		6	50	4	0.793	0.150	5.30
			50	4	0.440	0.183	2.41
			50	4	0.765	0.304	2.52
		7	50	4	0.614	0.089	6.86
			50	4	0.372	0.085	4.39
			50	4	0.581	0.122	4.78
A	Nickel sulfate	3	1	4	0.303	0.221	1.37
			3	4	0.424		1.92
			10	4	0.570		2.58
		4	1	4	0.431	0.210	2.05
			3	4	0.420		2.00
			10	4	0.952		4.53
		7	1	4	0.273	0.145	1.88
			3	4	0.386		2.66
			10	4	0.385		2.66
B	Isopropanol	1	10	4	0.350	0.158	2.22
			25	4	0.120		0.76
			50	4	0.145		0.92
		3	10	4	0.261	0.266	0.98
			25	4	0.227		0.85
			50	4	0.199		0.75
		4	10	4	0.240	0.241	1.00
			25	4	0.292		1.21
			50	4	0.380		1.58
		5	10	4	0.052	0.055	0.94
			25	4	0.038		0.69
			50	4	0.040		0.71
		6	10	3	0.516	0.253	2.04
			25	3	0.283		1.12
			50	3	0.383		1.51
		7	10	4	0.058	0.120	0.48
			25	4	0.115		0.95
			50	4	0.121		1.01
C	Eugenol	2	10	4	0.226	0.173	1.31
			25	4	0.422		2.45
			50	4	0.546		3.17
		6	10	4	0.306	0.210	1.46
			25	4	0.573		2.73
			50	4	0.667		3.18
		7	10	4	0.359	0.123	2.92
			25	4	0.514		4.18
			50	4	0.870		7.08
D	Cinnamic aldehyde	2	1	4	0.196	0.178	1.10
			3	4	0.397		2.23
			10	4	0.600		3.37
		4	1	4	0.426	0.271	1.57
			3	4	0.796		2.94
			10	4	0.947		3.49
		5	1	4	0.171	0.150	1.14
			3	4	0.315		2.10
			10	4	0.617		4.11
E	2,4-Dinitrochlorobenzene	1	0.1	4	0.674	0.302	2.23
			0.3	4	1.110		3.67
			1.0	4	1.298		4.30
		2	0.1	4	1.137	0.178	6.39
			0.3	4	1.162		6.52
			1.0	4	1.490		8.36

Table 3. *Continued*

Code	Chemical name	Lab. no.	Dose	*n*	Mean absorbance for chemical	Mean absorbance for vehicle	SI
		3	0.1	4	0.941	0.220	4.27
			0.3	4	1.378		6.25
			1.0	4	1.319		5.99
		4	0.1	4	1.005	0.271	3.71
			0.3	4	1.434		5.29
			1.0	4	1.490		5.50
		5	0.1	4	2.243	0.150	14.94
			0.3	4	2.819		18.78
			1.0	4	2.540		16.93
		6	0.1	4	0.711	0.210	3.38
			0.3	4	0.944		4.50
			1.0	4	1.014		4.83
		7	0.1	4	0.705	0.123	5.73
			0.3	4	1.509		12.28
			1.0	4	1.593		12.96
F	Glutaraldehyde solution	1	0.1	4	0.188	0.107	1.76
			0.3	4	0.257		2.40
			1.0	4	0.400		3.73
		5	0.1	4	0.395	0.053	7.44
			0.3	4	0.689		12.98
			1.0	4	1.525		28.73
		6	0.1	4	0.162	0.163	0.99
			0.3	4	0.308		1.89
			1.0	4	0.367		2.25
G	Methyl salicylate	1	10	4	0.431	0.302	1.43
			25	4	0.417		1.38
			50	4	0.381		1.26
		2	10	4	0.192	0.173	1.11
			25	4	0.201		1.16
			50	4	0.248		1.44
		3	10	4	0.242	0.220	1.10
			25	4	0.267		1.21
			50	4	0.309		1.40
H	Hexylcinnamic aldehyde	1	10	4	0.248	0.158	1.57
			25	4	0.412		2.61
			50	4	0.537		3.41
		3	10	4	0.491	0.266	1.20
			25	4	0.625		2.06
			50	4	0.804		2.87
		4	10	4	0.491	0.241	2.04
			25	4	0.625		2.59
			50	4	0.804		3.34
		5	10	4	0.291	0.055	5.25
			25	4	0.474		8.57
			50	4	0.746		13.48
		6	10	4	0.450	0.253	1.78
			25	4	0.727		2.87
			50	4	0.827		3.27
		7	10	4	0.192	0.120	1.59
			25	4	0.366		3.04
			50	4	0.462		3.84
I	Lactic acid	3	10	4	0.241	0.221	1.09
			25	4	0.365		1.66
			50	4	0.397		1.80
		4	10	4	0.359	0.210	1.71
			25	4	0.397		1.89
			50	4	0.343		1.63
		7	10	4	0.175	0.145	1.21
			25	4	0.313		2.16
			50	4	0.367		2.53
J	Formaldehyde solution	1	1	4	0.330	0.107	3.08
			3	4	0.471		4.40
			10	4	0.191		1.78
		5	1	4	0.225	0.053	4.25
			3	4	0.088		1.67
			10	4	0.883		16.64
		6	1	4	0.261	0.163	1.60
			3	4	0.293		1.80
			10	4	0.321		1.97

Journal of AppliedToxicology

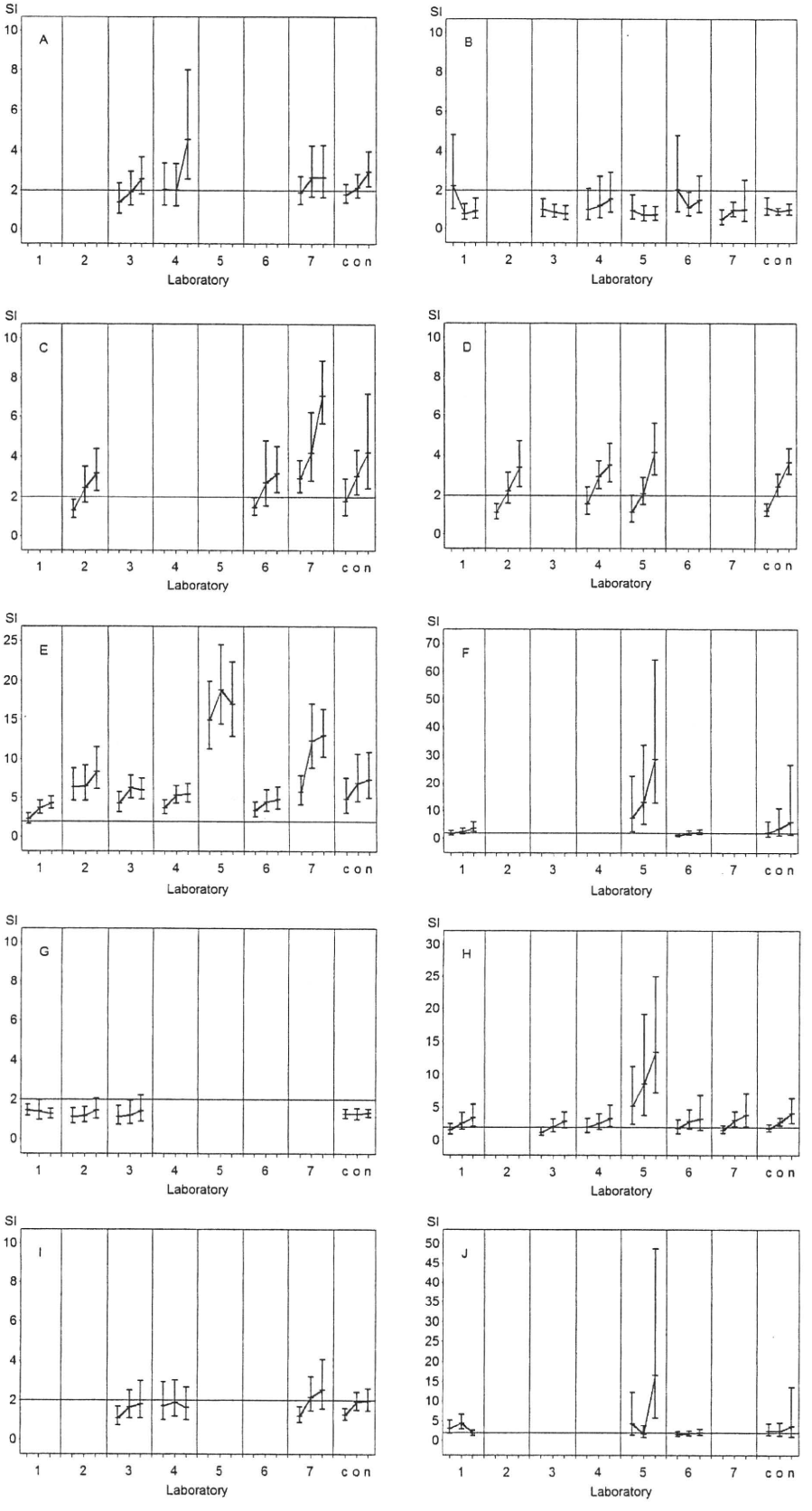

Figure 3. Dose–response relationships of the SI values with 95% confidence intervals for each chemical analyzed at the all laboratories in phase II validation study.
WA indicates the weighted average of the SI values obtained by meta-analysis using the random effect model in this study.