Dean et al. (2001) noted that "each SI should include a measure of variability that takes into account the interanimal variability in both the dosed and the control groups"; however, they did not describe the approach available for the calculation. As an alternative to the ignorance approach, the variance of SI may be calculated by approximation; this is known as the delta method. The purpose of this article is (1) to derive a mathematical formula for calculating the variation in SI by using the delta method and (2) to compare the variances obtained by the ignorance approach and the delta method and evaluate which of these approaches is appropriate for use in statistical analyses based on SI.

2. Materials and Methods

2.1 Variances of SI

Let Mean(i) be the mean DPM/mouse for the i^{th} group, and let SE(i) be the standard error (SE) of this value for the i^{th} group; i indexes the chemical-treated group (Y) and the vehicle control group (X). Thus, it follows that SI = Mean(Y)/ Mean(X).

When the variance is calculated using the ignorance approach, the variance of SI, i.e., $\mathrm{Var(SI)}_{IA}$, can be expressed as follows:

$$\mathrm{Var(SI)}_{IA} = \frac{\mathrm{SE(Y)}^2}{\mathrm{Mean(X)}^2} \quad (1)$$

As described above, this approach does not reflect the variation in the vehicle control group. To overcome this limitation, we propose the use of another approach in which an approximation known as the delta method is applied twice. In other words, after the variance of the log-transformed SI, i.e., $\mathrm{Var(ln\,SI)}_{DLT}$, is estimated using the delta method, the variance of SI is estimated by reapplying the delta method to the log-transformed SI. We refer to the variance obtained in this case as $\mathrm{Var(SI)}_{DLT}$, and it can be expressed as follows:

$$\mathrm{Var(SI)}_{DLT} = (\mathrm{SI})^2 \times \mathrm{Var(ln\,SI)} \quad (2)$$

where

$$\mathrm{Var(ln\,SI)} = \frac{\mathrm{SE(Y)}^2}{\mathrm{Mean(Y)}^2} + \frac{\mathrm{SE(X)}^2}{\mathrm{Mean(X)}^2} \quad (3)$$

It should be noted that to calculate these variances, only the means and standard errors of the DPM/mouse values within each group are required, not each DPM/mouse value.

2.2 Evaluation

Based on a simple calculation, the following relationship between equations (1) and (2) can be demonstrated:

$$\mathrm{Var(SI)}_{DLT} - \mathrm{Var(SI)}_{IA} = (\mathrm{SI})^2 \times \frac{\mathrm{SE(X)}^2}{\mathrm{Mean(X)}^2} \quad (4)$$

This equation states that $\mathrm{Var(SI)}_{DLT}$ is always greater than $\mathrm{Var(SI)}_{IA}$ and that the difference between these variances increases with increase in the SI. Therefore, it is predicted that for a high SI value corresponding to a chemical providing severe stimulation, the widths of the confidence interval differ markedly and may occasionally lead to misinterpretation. Since $\mathrm{Var(SI)}_{IA}$ is obtained by ignoring the variation in the vehicle control group and $\mathrm{Var(SI)}_{DLT}$, by using an approximation in the calculation, both methods may be biased. Further, it is unclear whether these approaches are suitable for practical use because the true variance is unknown. We then conducted a simulation study in order to examine the magnitude of the difference between the variance values obtained by both methods under several conditions and to determine the approach that is appropriate for practical use. In this investigation, we obtained the true variance of the SI by a Monte Carlo simulation. In other words, we regarded the sample variance of the SIs calculated using the random values generated for 10,000 replications as a true variance of SI. In the simulation, random numbers for the mean DPM/mouse values were generated for the chemical-treated group (Y) and vehicle control group (X), and the SIs were then calculated. The mean DPM/mouse values were assumed to follow a normal distribution truncated at less than 20 with a mean of 200 and standard deviations of 10, 25, and 50; the mean DPM/mouse follows a normal distribution with a mean of 200, 600, 1,000, 2,000, and 10,000 and standard deviations of 10, 50, and 100. Under these conditions, the SI values were 1, 3, 5, 10, and 50. The reason for the use of truncated normal distribution in the vehicle control group was that the SI is very sensitive to lower denominator values, and it was considered that in practicality, the mean DPM/mouse values below 20 cannot be regarded to indicate a successful examination.

We also examined the extent of the difference between the 2 variances by using published real data regarding the LLNA. Several inter-laboratory studies on the LLNA have been conducted, wherein the means and standard errors of the DPM/mouse values in experiments on different chemicals have been described (Basketter et al., 1991; Kimber et al., 1991a, 1995b, 1998c; Loveless et al., 1996; Scholes et al., 1992). Loveless et al. (1996) reported the results of experiments conducted in 5 laboratories, wherein 7 chemicals were tested by using the LLNA; however, they reported the standard errors for only 3 laboratories. We used their published data regarding the values obtained for 3 different chemicals (dinitrochlorobenzene, isoeugenol, and para-aminobenzoic acid (PABA)) for our subsequent investigations (Table 1).

3 Results

3.1 Monte Carlo simulation

Table 2 shows the calculated variances, in which $Var(SI)_{SIM}$ was obtained from the Monte Carlo simulation and $Var(SI)_{IA}$ and $Var(SI)_{DLT}$ were calculated based on the values obtained under different simulation conditions. Since $Var(SI)_{IA}$ is estimated solely based on Mean(X) and SE(Y), the variance obtained under various conditions remains the same. When SE(X) = 10 and SI = 1, both $Var(SI)_{IA}$ and $Var(SI)_{DLT}$ were almost equal and were less biased. However, these values differed under other conditions. In particular, they differed considerably in the case of high SI values, and $Var(SI)_{IA}$ was extremely biased toward underestimation. When SE(X) = 10, $Var(SI)_{DLT}$ was slightly biased, whereas for high SE(X) values, it tended to be underestimated.

3.2 Numerical examination

Table 3 shows the SIs and variances obtained using the different approaches in each laboratory and for each dose of the 3 chemicals listed in Table 1. For higher SIs, the difference between the values calculated using the 2 approaches tended to be greater. For example, in the case of 0.25% dinitrochlorobenzene examined in laboratory 3, the $Var(SI)_{DLT}$ value was 112.59, while the $Var(SI)_{IA}$ value was 4.47. This is due to the first condition of equation (4).

Table 1. Mean and standard error of the DPM/mouse values for each group in 3 laboratories. These data were reported by Loveless et al. (1996). It is reported that the number of mice / group was 4 or 5.

	Lab. 1		Lab. 2		Lab. 3	
Concentration (%)	Mean	SE	Mean	SE	Mean	SE
Dinitrochlorobenzene						
Solvent	287	64	163	22	26	11
0.010	444	67	163	18	65	12
0.025	529	179	194	27	76	28
0.050	675	119	455	93	82	13
0.100	2550	349	2092	316	184	23
0.250	10953	493	12814	1675	639	55
Isoeugenol						
Solvent	251	22	313	57	43	12
0.25	729	105	228	39	53	11
0.50	435	112	230	37	74	77
1.00	584	40	272	10	112	16
2.50	953	145	649	133	184	35
5.00	1718	259	2242	487	479	96
pABA						
Solvent	90	13	139	18	59	16
0.5	101	14	223	33	67	10
1.0	98	10	116	11	37	10
2.5	104	16	121	18	39	11
5.0	100	23	91	5	46	10
10.0	86	15	97	3	35	7

SE, Standard error.

Table 2. Theoretical examination of 3 variances obtained under several conditions.
The mean DPM/mouse value in the chemical-treated group was set at 200. $Var(SI)_{SIM}$ was obtained in the Monte Carlo simulation, and $Var(SI)_{IA}$ and $Var(SI)_{DLT}$ were calculated from the values obtained in the simulation.

		SE(X)								
		10			25			50		
SE(Y)	SI	$Var(SI)_{SIM}$	$Var(SI)_{IA}$	$Var(SI)_{DLT}$	$Var(SI)_{SIM}$	$Var(SI)_{IA}$	$Var(SI)_{DLT}$	$Var(SI)_{SIM}$	$Var(SI)_{IA}$	$Var(SI)_{DLT}$
10	1	0.005	0.003	0.005	0.021	0.003	0.018	0.138	0.003	0.065
50	1	0.066	0.063	0.065	0.083	0.063	0.078	0.219	0.063	0.125
100	1	0.253	0.250	0.253	0.281	0.250	0.266	0.467	0.250	0.313
10	3	0.025	0.003	0.025	0.164	0.003	0.143	1.244	0.003	0.565
50	3	0.087	0.063	0.085	0.227	0.063	0.203	1.376	0.063	0.625
100	3	0.275	0.250	0.273	0.421	0.250	0.391	1.530	0.250	0.813
10	5	0.066	0.003	0.065	0.449	0.003	0.393	3.376	0.003	1.565
50	5	0.127	0.063	0.125	0.517	0.063	0.453	3.594	0.063	1.625
100	5	0.316	0.250	0.313	0.704	0.250	0.641	3.633	0.250	1.813
10	10	0.256	0.003	0.253	1.794	0.003	1.565	13.191	0.003	6.253
50	10	0.318	0.063	0.313	1.840	0.063	1.625	13.832	0.063	6.313
100	10	0.509	0.250	0.500	2.049	0.250	1.813	14.070	0.250	6.500
10	50	6.442	0.003	6.253	44.352	0.003	39.065	338.721	0.003	156.253
50	50	6.415	0.063	6.313	45.240	0.063	39.125	346.652	0.063	156.313
100	50	6.626	0.250	6.500	44.519	0.250	39.313	343.594	0.250	156.500

$Var(SI)_{SIM}$, the variance of SI obtained from the simulation; $Var(SI)_{IA}$, the variance of SI obtained by the ignorance approach; $Var(SI)_{DLT}$, the variance of SI obtained by the delta method; SE(X), standard error for the DPM/mouse values in the vehicle control group; SE(Y), standard error for the DPM/mouse values in the chemical-treated group; and SI, Stimulation index.

Table 3. Numerical comparison of the variances obtained using the ignorance approach and the delta method approach.

Concentration (%)	Lab. 1			Lab. 2			Lab. 3		
	SI	$Var(SI)_{IA}$	$Var(SI)_{DLT}$	SI	$Var(SI)_{IA}$	$Var(SI)_{DLT}$	SI	$Var(SI)_{IA}$	$Var(SI)_{DLT}$
Dinitrochlorobenzene									
0.01	1.55	0.05	0.17	1.00	0.01	0.03	2.50	0.21	1.33
0.025	1.84	0.39	0.56	1.19	0.03	0.05	2.92	1.16	2.69
0.05	2.35	0.17	0.45	2.79	0.33	0.47	3.15	0.25	2.03
0.1	8.89	1.48	5.40	12.83	3.76	6.76	7.08	0.78	9.75
0.25	38.16	2.95	75.38	78.61	105.60	218.18	24.58	4.47	112.59
Isoeugenol									
0.25	2.90	0.17	0.24	0.73	0.02	0.03	1.23	0.07	0.18
0.5	1.73	0.20	0.22	0.73	0.01	0.03	1.72	3.21	3.44
1	2.33	0.03	0.07	0.87	0.00	0.03	2.60	0.14	0.67
2.5	3.80	0.33	0.44	2.07	0.18	0.32	4.28	0.66	2.09
5	6.84	1.06	1.42	7.16	2.42	4.12	11.14	4.98	14.65
pABA									
0.5	1.12	0.02	0.05	1.60	0.06	0.10	1.14	0.03	0.12
1	1.09	0.01	0.04	0.83	0.01	0.02	0.63	0.03	0.06
2.5	1.16	0.03	0.06	0.87	0.02	0.03	0.66	0.03	0.07
5	1.11	0.07	0.09	0.65	0.00	0.01	0.78	0.03	0.07
10	0.96	0.03	0.05	0.70	0.00	0.01	0.59	0.01	0.04

$Var(SI)_{IA}$, the variance of SI obtained using the ignorance approach; $Var(SI)_{DLT}$, the variance of SI obtained using the delta method; and SI, stimulation index.

4. Discussion

One of the merits of the LLNA when compared with other *in vivo* tests such as the guinea pig maximization test is that it reduces and refines animal use in identifying the hazards of skin-sensitizing chemicals (Basketter, 2005). Another important merit of this method is that it can quantitatively assess chemicals based on the SI. Despite this fact, variations in the SI have not been investigated sufficiently. In this study, we derived a formula for calculating the SI variance by using the delta method.

In our comparative investigation of the variances obtained using 2 different approaches, we observed that the value obtained using the ignorance approach, i.e., $\mathrm{Var}(\mathrm{SI})_{\mathrm{IA}}$, was extremely underestimated in the case of high SI values. In the LLNA, the SI cut-off value for judging whether the response to a chemical is positive or negative is usually set to be 3. It is possible that the dissimilarity between the 2 variances even around an SI value of 3 may be so great that the value obtained by the ignorance approach cannot be accepted. Therefore, to judge whether the response to a chemical is positive or negative, a statistical test based on the ignorance approach should not be used because it may yield an excessive number of false positive judgments. The delta method is a better approach than the ignorance approach.

Another important merit of using the variance of SI is with regard to the confidence interval. There exists a close link between the use of a confidence interval and a two-sided statistical test. Thus, the results of the statistical test can be inferred once the confidence interval has been calculated. Furthermore, presenting the confidence interval for the SI can directly reveal the SI precision (Gardner et al., 2000). Although we can simply obtain the 95% confidence interval as

$$\mathrm{SI} \pm 1.96 \times \sqrt{\mathrm{Var}(\mathrm{SI})} \qquad (5)$$

the lower limit value obtained by this equation may be an improbable value of less than 0. To eliminate this possibility, the confidence interval for the log-transformed SI can be exponentiated as follows:

$$\exp\left(\ln(\mathrm{SI}) \pm 1.96\sqrt{(\mathrm{Var}(\ln \mathrm{SI}))}\right) \qquad (6)$$

and $\mathrm{Var}(\ln \mathrm{SI})$ can be obtained from equation (3). This is another benefit of using the delta approach.

In conclusion, statistical analysis based on the ignorance approach is not acceptable because the variance thus obtained is severely biased toward underestimation. Instead, the delta method approach is recommended for practical use. However, it should be noted that this approach is also slightly biased in the case of large variations in the DPM/mouse values in the vehicle control group.

Acknowledgment

This work was supported by a Grant-in-Aid for Scientific Research (16200022).

References

Basketter, D. A., Casati, S., Gerberick, G. F., Griem, P., Philips, B., and Worth, A. (2005) Skin sensitisation, *Alternatives to Laboratory Animals*, 33, 83-103.

Basketter, D. A., Scholes, E. W., Kimber, I., Botham, P. A., Hilton, J., Miller, K., Robbins, M. C., Harrison, P. T. C., and Waite, S. J. (1991) Interlaboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data, *Toxicology Methods*, 1, 30-43.

Dean, J. H., Twerdok, L. E., Tice, R. R., Sailstad, D. M., Hattan, D. G., and Stokes, W. S. (2001). ICCVAM evaluation of the murine local lymph node assay. II. Conclusions and recommendations of an independent scientific peer review panel, *Regulatory Toxicology and Pharmacology*, 34, 258-273.

Gardner, M. J. and Altman, D. G. (2000) Confidence intervals rather than P values, in *Statistics with Confidence*, ed. by D. G. Altman, D. Machin, T. N. Bryant, and M. J. Gardner, pp. 15-27, BMJ Publishing, London.

Kimber, I., Hilton, J., Botham, P. A., Basketter, D. A., Scholes, E. W., Miller, K., Robbins, M. C., Harrison, P. T. C., Gray, T. J. B., and Waite, S. J. (1991) The murine local lymph node assay: results of an inter-laboratory trial, *Toxicology Letters*, 55, 203-213.

Kimber, I., Hilton, J., Dearman, R. J., Gerberick G. F., Ryan, C. A., Basketter, D. A., Scholes, E. W., Ladics, G. S., Loveless, S. E., House, R. V., and Guy, A. (1995) An international evaluation of the murine local lymph node assay and comparison of modified procedures, *Toxicology*, 103, 63-73.

Kimber, I., Hilton, J., Dearman, R. J., Gerberick, G. F.,

Ryan, C. A., Basketter, D. A., Lea, L., House, R. V., Ladics, G. S., Loveless, S. E., and Hastings, K. (1998) Assessment of the skin sensitization potential of topical medicaments using the local lymph node assay: an interlaboratory evaluation, *Journal of Toxicology and Environmental Health, Part A*, 53, 563-579.

Loveless, S. E., Ladics, G. S., Gerberick, G. F., Ryan, C. A., Basketter, D. A., Scholes, E. W., House, R. V., Hilton, J., Dearman, R. J., and Kimber, I. (1996) Further evaluation of the local lymph node assay in the final phase of an international collaborative trial, *Toxicology*, 108, 141-152.

OECD (2002) Organization for Economic Cooperation and Development – OECD Guidelines for Testing of Chemicals. No. 429: Skin Sensitization: The Local Lymph Node Assay.

Scholes, E. W., Basketter, D. A., Sarll, A. E., Kimber, I., Evans, C. D., Miller, K., Robbins, M. C., Harrison, P. T. C., and Waite, S. J. (1992) The local lymph node assay: results of a final inter-laboratory validation under field conditions, *Journal of Applied Toxicology*, 12, 217-222.

(Received: February 21, 2007 /
Accepted: September 30, 2007)

Corresponding author:
Takashi Omori, Ph.D.
Department of Biostatistics, Kyoto University,
Yoshida-Konoecho, Sakyo-ku, Kyoto 606-8501, Japan
Tel:+81-75-753-4482
Fax:+81-75-753-4487
E-mail: omori@pbh.med.kyoto-u.ac.jp

AATEX 14, Special Issue, 429-432
Proc. 6th World Congress on Alternatives & Animal Use in the Life Sciences
August 21-25, 2007, Tokyo, Japan

Validation studies on an alternative endpoint for the local lymph node assay (LLNA-DA): Importance of study management

Takashi Omori[1], Yoshiaki Ikarashi[2], Yukiko Kanazawa[3], Kenji Idehara[4], Hajime Kojima[2], Takashi Sozu[5], Kazunori Arima[6], Hirohiko Goto[7], Tomohiko Hanada[8], Taketo Inoda[9], Tadashi Kosaka[10], Eiji Maki[11], Takashi Morimoto[12], Shinsuke Shinoda[13], Naoki Shinoda[14], Masahiro Takeyoshi[15], Masashi Tanaka[16], Mamoru Uratani[17], Masahito Usami[18], Atsushi Yamanaka[19], Tomofumi Yoneda[20], Isao Yoshimura[21] and Atsuko Yuasa[22]

[1]Kyoto University SPH, [2]National Institute of Health Sciences, [3]Food and Drug Safety Center, [4]Daicel Chemical Industries Ltd., [5]Osaka University, [6]Taisho Pharmaceutical Co. Ltd., [7]Otsuka Pharmaceutical Co. Ltd., [8]Nippon Shinyaku Co. Ltd., [9]Nakano Seiyaku Co. Ltd., [10]Institute of Environmental Toxicology, [11]Biosafety Research Center, Foods, Drugs and Pesticides, [12]Sumitomo Chemical Co. Ltd., [13]Drug Safety Testing Center Co. Ltd., [14]Santen Pharmaceutical Co. Ltd., [15]Chemicals Evaluation and Research Institute, [16]Meiji Seika Kaisha Ltd., [17]Ishihara Sangyo Kaisha Ltd., [18]Hoyu Co. Ltd., [19]Pias Corporation, [20]Toaeiyo Ltd., [21]Tokyo University of Science and [22]Fuji Film Co. Ltd.

Corresponding author: Takashi Omori
Department of Biostatistics, Kyoto University School of Public Health
Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan
Phone: +(81)-75-753-4482, Fax: +(81)-75-753-4487, omori@pbh.med.kyoto-u.ac.jp

Abstract

We conducted 2 validation studies for a modified version of the local lymph node assay (LLNA), which was designated as the LLNA-DA. A total of 17 laboratories tested the validity of the assay by using 14 chemicals. Here, in addition to the experimental protocol, we prepared the study protocols describing the study purpose, the role of the participants, etc. Technology transfer was conducted by the developer of the assay. Prior to the studies, preliminary tests using only a positive control chemical were conducted to determine whether the experimental protocol prescribed for the assay was appropriate. A formatted data file was developed for data management. Fortunately, the results of these studies revealed small interlaboratory variations, and we believe that one of the factors that contributed to the successful results was the development of strategies and tools for study management at the planning stage itself. However, issues related to the management of validation studies have rarely been discussed. Strategies or tools developed for study management should be easily accessible and should be shared with researchers intending to conduct validation studies in the future.

Keywords: interlaboratory validation study, study management, protocol, technical transfer, data quality

Introduction

An interlaboratory validation study examines the reliability and relevance of a particular test method (Organization of Economic Co-operation and Development (OECD), 2005). It differs from a single laboratory study in that it involves many persons having different backgrounds and levels of experience. To minimize interlaboratory variations, it is necessary that all the participating researchers from each laboratory understand how to operate the test method and perform it accurately, according to the procedure specified for the study rather than the customary procedure used in their respective laboratories. Therefore, appropriate management is one of the challenges encountered in the success of an interlaboratory validation study.

The murine local lymph node assay (LLNA) has developed as an alternative to the guinea pig test for assessing skin sensitization. In this method, lymphocyte proliferation in the draining auricular lymph nodes is measured by the incorporation of radioactive molecules (OECD, 2002). Recently, several nonradioactive methods have been proposed. Daicel Chemical Industries Ltd. has developed a modified nonradioactive version of the LLNA that is based on the ATP content (Yamashita, 2005). Since this method was originated by Daicel Chemical Industries Ltd. and is based on the ATP content,

it is designated as the LLNA-DA. To evaluate the LLNA-DA, 2 validation studies were conducted by 23 researchers from 22 organizations. The first study examined the reliability and relevance of the method using 12 chemicals in 10 experimental laboratories. The second study examined the reliability of the method using 5 chemicals in 7 experimental laboratories.

Since these validation studies were conducted on a large scale, appropriate management was essential. Therefore, strategies and tools were developed for their management. Fortunately, the results of these studies on the LLNA-DA successfully revealed small interlaboratory variations and good relevance. We believe that one of the factors that contributed to the good results was the strategies and tools employed for managing the study. However, issues related to the management of validation studies have rarely been reported.

In this article, we report the strategies and tools that we developed for managing the LLNA-DA validation studies. First, we describe the 2 protocols used. Next, we discuss the seminar for technology transfer and the preliminary tests that were conducted. Further, we introduce the web folder that was developed for use, and we subsequently describe the formatted data file. Finally, we discuss the management of the validation studies and present our conclusion.

Two types of protocols

In commonly used dictionaries, the word "protocol" is defined as the plan for a medical treatment course or for a scientific experiment or as a predefined written procedure for designing and implementing experiments. In the context of clinical studies, its meaning is more specific. The word protocol describes a method to be used in a clinical trial or a medical research study. With regard to the purpose of a protocol in clinical studies, Collins (2001) states that "It describes in a clear and detailed manner how the trial is performed so that all investigators know the procedures. This is particularly important in multicenter trials where it can be difficult to ensure that all centers and investigators conduct the study properly." The difficulty encountered in multicenter trials that he states here is identical to that encountered in an interlaboratory validation study. Therefore, this type of a protocol that describes the method to be followed for performing various steps in a validation study should be required. On the other hand, the OECD guidance document 34 (OECD, 2005) defines a protocol as "the detailed, unambiguous step-by-step description of a test method that directs the laboratory as to how to perform the test method." In this case, the protocol pertains to the implementation of a test method but not to a validation study for the test method. Most biologists appear to be familiar with this definition, and without doubt, this type of protocol is also required in a validation study.

Therefore, 2 types of protocols were prepared for the validation study of the LLNA-DA. We designated the first document as the study protocol and the second one, as the experimental protocol. Fig. 1 shows the table of contents of the study protocol used for our study.

1. Introduction
2. Purpose of the study
3. Role of the researchers
4. Standard operating procedure for LLNA-DA
5. Time schedule
6. Participant organization
7. Chemicals tested
8. Chemical allocation
9. Preparation of animals, equipment, and materials
10. Expenditure
11. Technology transfer and preliminary test
12. Data management
13. Data analysis
14. Meeting held to discuss the results
15. Announcement of the results
16. Inquiries

Fig. 1. Table of contents of the protocol employed for the first study on the LLNA-DA.

Seminar for technology transfer and preliminary tests

Even if a well-documented experimental protocol is prepared, toxicologists from different laboratories may interpret the document differently. In order to determine their understanding of the experimental protocol and to explain the execution of the test method, a 1-day seminar for technology transfer was held by the LLNA-DA developer prior to each study. It was required to be attended by at least 1 toxicologist from each experimental laboratory.

To confirm that the experimental protocol was being adequately documented, a preliminary test employing only a positive control chemical, namely, hexylcinnamic aldehyde, was conducted prior to each study.

Fig. 2 (a) and (b) shows the results of the preliminary tests performed for each study. The plot illustrates the stimulation index (SI) value, which is the endpoint of interest in the LLNA-DA and is defined as an increase in the ATP content in the chemical-treated group relative to the vehicle control group, along with its 95% confidence intervals for all the laboratories. Only one experimental dose was used in the preliminary test for the first study; however, in order to assess the dose-response relationships, 2 different doses were used in the preliminary test for the second study. Based on these plots and the historical data obtained from Daicel

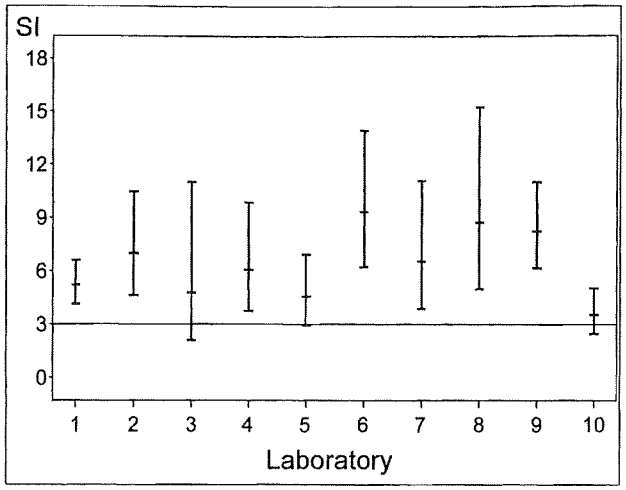

Fig. 2. (a) SI value with 95% confidence intervals obtained for the positive control chemical (25% hexylcinnamic aldehyde) in the preliminary test performed during the first study.

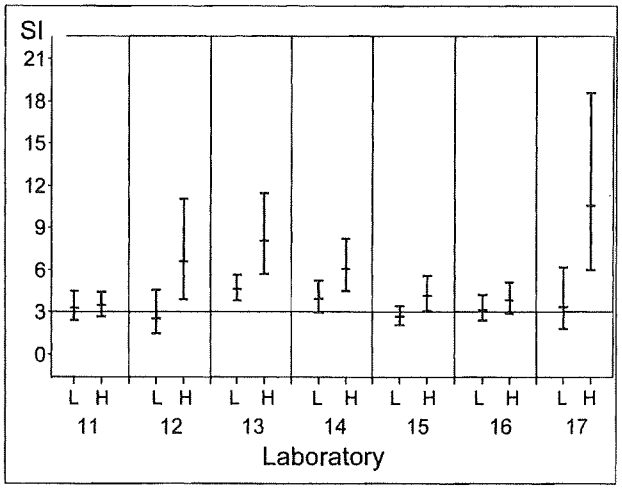

Fig. 2. (b) SI value with 95% confidence intervals obtained for the positive control chemical (10% (L) and 25% (H) hexylcinnamic aldehyde) in the preliminary test performed during the second study.

Chemical Industries Ltd., we discussed whether revisions were needed in the experimental protocol.

Use of a folder on a website

During a project, many documents related to a validation study are repeatedly revised to ensure that they reflect the opinions of each researcher. One of problems is that often important documents are lost or may fail to be updated. Therefore, all the researchers involved in the study are required to be well versed with the latest version of the documents.

To enable easy access to the latest version of the necessary documents pertaining to the validation studies, we used a commercial web tool, i.e., a folder on the website. By using this tool, all researchers could download the document via the internet onto any personal computer at their respective workplaces as and when required. Once a document was uploaded onto the web folder, it could be downloaded at any time. The web folder was set such that only the study manager was able to update the documents. When the study manager decided to upload or update a document, he accessed the web folder and uploaded the latest version of the document and then deleted the older version from the folder. Subsequently the study manager would then inform all the researches that the document had been updated. This rule was strictly followed throughout the study.

Formatted data file

To directly collect the raw data obtained from the experimental laboratories and to construct a database, an MS-Excel formatted file was prepared for entering the experimental data. One of the advantages of MS-EXCEL is that it is widely available, and many researchers can use it at their respective workplaces. Another advantage is that it has several useful functions. For example, it is possible to protect the data from being entered into an unintentional cell on the formatted file.

The empty formatted data file along with a document describing how it was to be used was distributed to the experimental laboratories prior to commencement of the experiment. Following data entry into the formatted data file, the file and the record that was maintained for the values observed during the experiment were collected from all the experimental laboratories. A biostatistician examined the values in both the file and the record. When needed, the toxicologist who carried out the LLNA-DA in the experimental laboratory was inquired about it. After resolving this issue, the biostatistician constructed a database on which all the data analysis was carried out. The purpose of constructing such a database is to ensure that the quality of the data is maintained.

Discussion

The OECD guidance document 34 (OECD, 2002) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) guidelines for the nomination and submission of new, revised, and alternative test methods (ICCVAM, 2003) are excellent documents that provide very useful information for conducting validation studies. However, both these documents focus on broad issues and are written from a more general viewpoint. On the other hand, here, we describe the management strategies and tools for a validation study from a more practical viewpoint, arising from discussions regarding some validation studies that have been conducted in Japan. In particular, some of the authors who were involved in the interlaboratory validation study for alternatives to the Draize eye irritation test, organized by the Japanese Society of Alternatives to

Animal Experiments (Ohno et al., 1998), participated. This study was conducted on a large scale and evaluated 16 cytotoxicity tests as alternative tests. The total number of experimental laboratories participating in the study was 16–24 per cytotoxicity test. Large interlaboratory variations were obtained for all the cytotoxicity tests, and it was very difficult to interpret the data and evaluate the cytotoxicity tests based on the study results because there were many instances of violation of rules that had been finalized prior to the study and misinterpretation of the experimental protocols (Omori, 1998). To clarify the purpose of the validation study, i.e., evaluating the interlaboratory variations under the experimental protocol, to transfer the experimental operations for the tests correctly, and to try to ensure data quality should have been considered from planning stage of the study. The study demonstrated that an interlaboratory validation study is a joint venture by researchers having different backgrounds and levels of experience. In other words, study management of the validation studies became a challenging issue.

We admit that the strategies and tools described here do not cover all the aspects of study management and that the strategy and tools for other validation studies should be developed by considering individual cases and various viewpoints. However, our strategies and tools proved to be efficient for at least 2 validation studies, and we believe that they could serve as a reference for researchers conducting validation studies for a test method in the future.

It is important to note that in addition to the processes described in the experimental protocol adopted for a test method, there are many factors that can contribute to the occurrence of large interlaboratory variations. In other words, it is possible that variations could arise in an established test method even if the experimental protocol is well defined and the test is conducted under Good Laboratory Practice conditions. To exclusively evaluate the test method described in the experimental protocol, attempts should be made to eliminate additional factors that could cause interlaboratory variations. Large interlaboratory variations would lead to unclear results from the study and would delay the development of a test method even in case of a well-defined method.

In conclusion, management implies all the activities that are necessary to achieve objectives continually and efficiently. To obtain scientifically valid and distinct results from a validation study, appropriate study management from the planning stage is critical. The knowledge base on the management of validation studies should be expanded and shared.

References

Collins, J. F. (2001) Protocol, In Biostatistics in Clinical Trials, ed. by P. Armitage and T. Colton, pp. 373–377, John Wiey & Sons, West Sussex.

ICCVAM. (2003) ICCVAM Guidelines for the Nomination and Submission of New, Revised, and Alternative Test Methods.

OECD. (2002) Organization for Economic Co-operation and Development—OECD Guidelines for Testing of Chemicals. No. 429: Skin sensitization: Local Lymph Node Assay.

OECD. (2005) Organization for Economic Co-operation and Development—OECD series on testing and assessment. No. 34: Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment.

Ohno, T., Asakura, M., Awago, T., Futamura, Y., Harihara, A., Hatao, M., Hattori, C., Hayasaka, A., Hayashi, M., Yahashi, T., Hirata, Z., Hori, H., Hoshi, H., Imai, K., Imazeki, I., Ishibashi, T., Itagaki, H., Iwata, T., Kakuma, M., Kaneda, S., Kato, I., Kato, M., Kawahatsu, T., Kawakami, A., Kazama, A., Kido, A., Kimura, S., Kitazawa, M., Kogiso, S., Kojima, H., Kotani, M., Kuramochi, M., Maki, D., Matsuda, M., Mitsuoka, C., Miyazaki, S., Mizuno, F., Mori, M., Morimoto, K., Moriysu, M., Nakajima, K., Nakajima, M., Nakamura, M., Nakamura, M., Nakano, N., Nakamura, S., Inegami, A., Nishino, M., Nishitomi, T., Ohkoshi, T., Okamoto, Y., Omori, T., Ono, H., Ono, M., Osanai, Y., Saijo, K., Sano, Y., Saotome, K., Sasaki, K., Sasaki, T., Sato, H., Sato, S., Shimada, H., Shimogo, S., Shimono, K., Shionoya, H., Sugawara, H., Sugiki, Y., Sugimoto, S., Sugimoto, S., Suzuki, J., Takagaki, K., Takahashi, K., Takizawa, M., Tamaki, C., Tanaka, N., Taniya, J., Teramoto, N., Torishima, H., Tsuchiya, T., Uejima, M., Ueno, H., Ugai, Y., Wada, S., Wakuri, S., Wang, X., Watanabe, I., Watanabe, M., Yajima, S., Yamagata, Y., Yamaguchi, Y., Yamakita, O., Yamamoto, R., Yoshida, M., Yoshimura, I., Yuhki, K., and Yukiyama, S. (1998) Validation study on five cytotoxicity assays by JSSSE—I. Overview of the study and analyses of variations of ED_{50} Values. *Alternatives to Animal Testing and Experimentation* **5**, 1–38.

Omori, T., Saijo, K., Kato, M., Itagaki, H., Hayashi, M., Miyazaki, S., Ohno, T., Sugawara, H., Teramoto, N., Tanaka, N., Wakuri, S., and Yoshimura, I. (1998) Validation study on five cytotoxicity assays by JSAAE—III. Quality of collected data files. *Alternatives to Animal Testing and Experimentation* **5**, 59–73.

Yamashita, K., Idehara, K., Fukuda, N., Yamagishi, G., and Kawada, N. (2005) Development of a modified local lymph node assay using ATP measurement as an endpoint. *Alternatives to Animal Testing and Experimentation* **11**, 136–144.

Journal of Pharmacological and Toxicological Methods 58 (2008) 11–26

Contents lists available at ScienceDirect

Journal of Pharmacological and Toxicological Methods

journal homepage: www.elsevier.com/locate/jpharmtox

Original article

Interlaboratory validation of the modified murine local lymph node assay based on adenosine triphosphate measurement

Takashi Omori [a,*], Kenji Idehara [b,*], Hajime Kojima [c], Takashi Sozu [d], Kazunori Arima [e], Hirohiko Goto [f], Tomohiko Hanada [g], Yoshiaki Ikarashi [c], Taketo Inoda [h], Yukiko Kanazawa [i], Tadashi Kosaka [j], Eiji Maki [k], Takashi Morimoto [l], Shinsuke Shinoda [m], Naoki Shinoda [n], Masahiro Takeyoshi [o], Masashi Tanaka [p], Mamoru Uratani [q], Masahito Usami [r], Atsushi Yamanaka [s], Tomofumi Yoneda [t], Isao Yoshimura [u], Atsuko Yuasa [v]

[a] *Kyoto University School of Public Health, Japan*
[b] *Daicel Chemical Industries Ltd., Japan*
[c] *National Institute of Health Sciences, Japan*
[d] *Osaka University, Japan*
[e] *Taisho Pharmaceutical Co. Ltd., Japan*
[f] *Otsuka Pharmaceutical Co. Ltd., Japan*
[g] *Nippon Shinyaku Co. Ltd., Japan*
[h] *Nakano Seiyaku Co. Ltd., Japan*
[i] *Food and Drug Safety Center, Japan*
[j] *Institute of Environmental Toxicology, Japan*
[k] *Biosafety Research Center, Foods, Drugs and Pesticides, Japan*
[l] *Sumitomo Chemical Co. Ltd., Japan*
[m] *Drug Safety Testing Center Co. Ltd., Japan*
[n] *Santen Pharmaceutical Co. Ltd., Japan*
[o] *Chemicals Evaluation and Research Institute, Japan*
[p] *Meiji Seika Kaisha Ltd., Japan*
[q] *Ishihara Sangyo Kaisha Ltd., Japan*
[r] *Hoyu Co. Ltd., Japan*
[s] *Pias Corporation, Japan*
[t] *Toa Eiyo Ltd., Japan*
[u] *Tokyo University of Science, Japan*
[v] *Fuji Film Co. Ltd., Japan*

ARTICLE INFO

Article history:
Received 27 March 2008
Accepted 7 May 2008

Keywords:
Adenosine triphosphate
Interlaboratory validation
Local lymph node assay based on adenosine triphosphate content
Local lymph node assay
Skin sensitization
Methods

ABSTRACT

Introduction: 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 drugs and chemicals. Daicel Chemical Industries Ltd. has developed a modified LLNA based on the adenosine triphosphate (ATP) content (LLNA-DA). We conducted 2 interlaboratory validation studies to evaluate the reliability and relevance of LLNA-DA. **Methods:** The experiment involved 17 laboratories, wherein 14 chemicals were examined under blinded conditions. In the first study, 3 chemicals were examined in 10 laboratories and the remaining 9 were examined in 3 laboratories. In the second study, 1 chemical was examined in 7 laboratories and the remaining 4 chemicals were examined in 4 laboratories. The data were expressed as the ATP content for each chemical-treated group, and the stimulation index (SI) for each chemical-treated group was determined as the increase in the ATP content relative to the concurrent vehicle control group. An SI of 3 was set as the cut-off value for exhibiting skin sensitization activity. **Results:** The results of the first study obtained in the experiments conducted for the 3 chemicals that were examined in all the 10 laboratories and for 5 of the remaining 9 chemicals were sufficiently consistent with small variations in their SI values. The sensitivity, specificity, and accuracy of LLNA-DA against those of GPMT/BT were 7/8 (87.5%), 3/3 (100%), and 10/11 (90.9%), respectively. In the second study, all the 5 chemicals studied demonstrated acceptably small interlaboratory variations. **Discussion:** In the first study, a large variation was observed for 2 chemicals; in the second study, this variation was small. It was attributed to the

* Corresponding authors. Omori is to be contacted at Department of Biostatistics, Kyoto University School of Public Health, Yoshida Konoe-cho, Sakyo-ku, Kyoto 606-8501, Japan. Tel.: +81 75 753 4482; fax: +81 75 753 4487. Idehara, Daicel Chemical Industries, Ltd., 1239 Shinzaike, Aboshi-ku, Himeji, Hyogo 671-1283, Japan. Tel.: +81 79 274 4096; fax: +81 79 274 5831.
E-mail addresses: omori@pbh.med.kyoto-u.ac.jp (T. Omori), kn_idehara@daicel.co.jp (K. Idehara).

1056-8719/$ - see front matter
doi:10.1016/j.vascn.2008.05.001

application of dimethylsulfoxide as the solvent for the metallic salts. In conclusion, these 2 studies provide good evidence for the reliability of the LLNA-DA.

1. Introduction

Skin sensitization (allergic contact dermatitis (ACD)) is an immunologically mediated cutaneous reaction to a drug or chemical. It is known that detecting and evaluating the immune-based adverse effects that are collectively referred to as hypersensitivity reactions is a very difficult task, 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).

The local lymph node assay (LLNA) employs a mouse model for assessing the relative sensitization potential; it 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, LLNA was developed on the basis of a mechanistic understanding of immune-based contact dermatitis (Hastings, 2001). In addition, this method also offers important animal welfare benefits. The use of LLNA has been successfully validated by several studies (Basketter et al., 2002; Basketter, Gerberick, Kimber, & Loveless, 1996; Basketter & Scholes, 1992; Gerberick, Ryan, Kimber, Dearman, & Basketter, 2000; Haneke, Tice, Carson, Margolin, & Stokes, 2001). Recently, it has been recommended that this method be formally adopted by the Organization for Economic Co-operation and Development (OECD), according to the guidelines for testing chemicals 406 and 429 (OECD, 1992, 2002), and that it be accepted by the EU and US as a suitable method for classifying the skin sensitizing ability of chemicals (Basketter, Casati, Gerberick, Griem, Philips, & Worth, 2005; Dean, Twerdok, Tice, Sailstad, Hattan, & Stokes, 2001; Sailstad, Hattan, Hill, & Stokes, 2001). The LLNA was specifically designed to identify contact allergens. The assay was 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, De Burlet, Virat, & Nauman, 2003). Furthermore, presently, 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 [^{3}H]-methyl thymidine to measure lymphocyte proliferation; this hinders it use, particularly in Japan, because being a radioisotope (RI)-based method, it requires special facilities. Several authors have been conducting investigations for the development of an alternative non-RI method for performing LLNA (Dearman, Hilton, Basketter, & Kimber,, 1999; Ehling et al., 2005a, 2005b; Hatao, Hariya, Katsumura, & Kato, 1995; Lee, Park, Park, Kim, & Oh, 2002; Takeyoshi, Yamasaki, Yakabe, Takatsuki, & Kimber, 2001).

Daicel Chemical Industries Ltd. proposed a modification of LLNA, which involves the measurement of the adenosine triphosphate (ATP) content instead of [^{3}H]-methyl thymidine incorporation for assessing lymphocyte proliferation (Idehara, Yamagishi, Yamashita, & Ito, in press; Yamashita, Idehara, Fukuda, Yamagishi, & Kawada, 2005). This modified assay method is designated as the LLNA modified by Daicel, based on the ATP content (LLNA-DA).

Although LLNA-DA essentially involves the same procedure as LLNA, the evidence available is insufficient for validating the assay method through interlaboratory evaluation. Therefore, we conducted 2 interlaboratory validation studies for LLNA-DA.

In the first study, 2 metallic salts—cobalt chloride and nickel sulfate—dissolved in dimethylsulfoxide (DMSO) produced inconsistent results across the laboratories. We assumed that the inconsistency factor would be due to one of the following 2 reasons: (1) DMSO was used as the vehicle in the control group for the 2 metallic salts, and DMSO application in mice is difficult as compared with acetone–olive oil (AOO) or acetone (ACE) application or (2) LLNA-DA is unsuitable for use with metallic salts, and both the chemicals used were metallic salts. Therefore, a second study employing additional metallic salt with DMSO was planned in order to ascertain the hypothesis.

The primary objectives of the first study were (1) to evaluate the extent of interlaboratory variation with regard to LLNA-DA and (2) to ascertain whether the results of LLNA-DA are comparable with those of LLNA. The primary objective of the second study was to examine the reliability of the LLNA-DA method when metallic salts were tested with DMSO.

2. Methods

2.1. Organization

This study was organized by researchers belonging to the committee for the validation of the assay. The research team comprised

Table 1(a)
Selected chemicals with their corresponding vehicles, the referenced results of LLNA and GPMT/BT, and the allocation of chemicals for the LLNA-DA experiments in the first study

Chemical	CASRN[a]	Vehicle[b]	LLNA	GPMT/BT[c]	Laboratory[d]									
					1	2	3	4	5	6	7	8	9	10
A: 2,4-Dinitrochlorobenzene	97-00-7	AOO	+	+	□	□	□	▯	○	△	▯	▯	△	○
B: Hexyl cinnamic aldehyde	101-86-0	AOO	+	+	○	○	△	△	△	▯	△	○	○	△
C: 3-Aminophenol	591-27-5	AOO	+	+nonstd	□		○					▯		
D: Glutaraldehyde	111-30-8	ACE	+		△	△			▯					
E: Cobalt chloride	7646-79-9	DMSO	+	+				○		○		△		
F: Isoeugenol	97-54-1	AOO	+	+					○				△	
G: Formaldehyde	50-00-0	ACE	+	+	△	△			▯					
H: Dimethyl isophthalate	1469-93-4	AOO	–	–	▯		▯				▯			
I: Isopropanol	67-63-0	AOO	–	–	○	○	△	△	△	▯	△	○	○	△
J: Nickel sulfate	10101-97-0	DMSO	–	+				○		○		△		
K: Abietic acid	514-10-3	AOO	+	+		▯				△	○			
L: Methyl salicylate	119-36-8	AOO	–	–			○				○			○

[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; "nonstd" indicates a nonstandard animal that was not tested for chemical G.
[d] Allocated pairs for the LLNA-DA experiments in a laboratory; ○, experiment 1; △, experiment 2; ▯, experiment 3.

Table 1(b)
Selected chemicals with their corresponding vehicles, the referenced results of LLNA and GPMT/BT, and the allocation of chemicals in the second study

Chemical	CASRN[a]	Vehicle[b]	LLNA	GPMT/BT[c]	Laboratory[d]						
					11	12	13	14	15	16	17
B: Hexyl cinnamic aldehyde	101-86-0	AOO	+	+	○	○	○	○	○	○	○
E: Cobalt chloride	7657-79-9	DMSO	+	+			△	△			△
J: Nickel sulfate	10101-97-0	DMSO	−	+		△		△		△	
M: Lactic acid	598-82-3	DMSO	−	−	△		△		△	△	
N: Potassium dichromate	7778-50-9	DMSO	+	+	△	△			△		△

[a] The Chemical Abstract Services Registry Number.
[b] ACE, acetone; AOO, acetone–olive oil; DMSO, dimethylsulfoxide.
[c] Judgment based on guinea pig maximization test or Buehler test.
[d] Allocated pairs for an experiment in a laboratory; ○, experiment 1; △, experiment 2; |, experiment 3.

representatives from each experimental laboratory, toxicologists as the chemical selectors and as distributors of the chemicals and materials, biostatisticians, and the study manager. All the experimentations were performed by the toxicologists of the experimental laboratories. In the first study, participation was limited to 10 experimental laboratories with sufficient experience in the use of the LLNA and/or its modifications; however, this was not a limiting factor in the second study, in which 7 additional experimental laboratories were included. A total of 17 different experimental laboratories participated in these 2 studies.

Research teams of all the experimental laboratories obtained ethical approval for each standard operational procedure conducted in their laboratories.

2.2. Technology transfer

A 1-day technology-transfer seminar was held by the LLNA-DA developer for each study, which was attended by at least 1 toxicologist from each experimental laboratory. Participants learned the method of conducting the assay according to the standard protocol. In addition, in the second study, the operation of LLNA-DA with DMSO was also included in the seminar (Omori et al., 2008).

2.3. Preliminary tests

Prior to each study, a preliminary test was conducted by researchers from all the experimental laboratories, who used only the positive control chemical, namely, 25% hexyl cinnamic aldehyde. The purpose of these preliminary tests was to ascertain whether the standard protocol was being documented sufficiently and to confirm the sensitivity of LLNA-DA (Omori et al., 2008).

The results of both preliminary tests revealed that the standard protocol was essentially valid and required few modifications.

2.4. 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 & Scholes, 1992; Basketter, Gerberick, & Kimber, 1998; Basketter, Lea, Cooper, et al., 1999; Basketter, Lea, Dickens, 1999; Basketter, Blaikie, Dearman, Kimber, Ryan, Gerberick, et al., 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 3 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 the first study, 12 of the 20 candidate chemicals were selected and classified as strong, mild, or weak sensitizers or non-sensitizers on the basis of LLNA. In order to reduce the number of animals used, pairs comprising groups treated with 2 or 3 chemicals and the same vehicle control group were employed; in other words, in each laboratory, 2 or 3 chemicals were simultaneously tested with 1 negative control and 1 positive control for every experiment. Of the 12 chemicals, 3 were dispatched to all the 10 participating experimental laboratories, and the remaining 9 were randomly allocated to the laboratories by a biostatistician and dispatched to each of the 3 experimental laboratories.

In the second study, 5 of the 20 candidate chemicals were selected. To determine whether the results from the 7 new laboratories would be similar to those obtained in the first study, the chemical selectors chose a single chemical that had been tested by all the 10 laboratories in the first study. The remaining 4 chemicals selected by the chemical selectors comprised 3 metallic salts—cobalt chloride, nickel sulfate, and potassium dichromate—and lactic acid with DMSO as the vehicle control. Pairs comprising groups treated with 2 of the 4 chemicals and

Table 2(a)
Body weight (g) [day 1]

Laboratory	*n*	Mean	SD	Min	Med	Max
1	120	22.0	1.5	19.3	21.8	27.1
2	108	22.5	1.3	19.4	22.6	25.0
3	108	22.0	1.2	18.2	22.0	24.8
4	108	22.7	1.4	20.0	22.5	26.7
5	108	21.6	1.1	19.1	21.6	24.4
6	108	21.7	1.4	19.3	21.7	24.9
7	108	22.8	1.4	18.5	22.8	25.9
8	108	23.4	1.5	20.5	23.3	28.6
9	72	23.0	1.2	20.1	22.9	26.5
10	72	22.6	1.4	19.8	22.5	25.8
11	96	22.9	1.3	19.9	22.9	26.5
12	60	21.6	1.0	18.8	21.7	24.1
13	60	22.2	1.1	19.5	22.1	24.8
14	60	21.8	1.5	18.7	21.8	24.3
15	60	22.5	1.1	20.0	22.5	25.2
16	60	22.3	1.5	18.8	22.6	25.5
17	60	22.1	1.4	19.5	22.3	26.4

Table 2(b)
Body weight (g) [day 8]

Laboratory	*n*	Mean	SD	Min	Med	Max
1	120	22.1	1.5	19.0	22.0	26.1
2	108	23.4	1.4	20.6	23.3	26.7
3	108	23.2	1.4	19.8	23.2	26.6
4	104	23.4	1.4	20.4	23.3	27.1
5	108	23.0	1.3	20.1	23.0	25.8
6	108	22.2	1.4	19.2	22.2	25.6
7	108	23.0	1.5	17.1	23.0	26.0
8	108	23.9	1.8	20.1	24.0	29.2
9	72	23.9	1.3	20.9	23.9	27.0
10	72	23.3	1.3	20.7	23.3	26.8
11	96	23.4	1.3	21.1	23.3	27.1
12	60	23.1	1.2	20.4	23.2	26.5
13	60	22.9	1.3	20.2	22.7	26.2
14	59	22.3	1.9	16.3	22.4	25.9
15	60	23.8	1.3	21.3	23.6	26.6
16	60	23.3	1.6	19.1	23.4	27.0
17	60	23.1	1.4	19.7	23.3	26.7

Table 3(a)
Mean and SD for the ATP content and SI values obtained in all the laboratories in the first study

Vehicle/ concentration	1		2		3		4		5		6		7		8		9		10	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
A: 2,4-Dinitrochlorobenzene																				
AOO	27,188±10,027	–	26,159±2157	–	35,610±7212	–	42,866±9956	–	11,899±7366	–	13,910±3921	–	22,466±3515	–	20,576±5546	–	26,842±9515	–	53,350±14,893	–
0.03%	77,305±25,181	2.8	60,843±19,746	2.3	80,548±34,265	2.3	127,990±23,651	3.0	18,107±3203	1.5	38,247±10,833	2.7	86,083±21,219	3.8	49,730±22,738	2.4	75,290±20,086	2.8	62,000±23,941	1.2
0.10%	147,161±32,102	5.4	70,451±26,337	2.7	150,579±23,446	4.2	210,206±57,119	4.9	45,691±21,305	3.8	59,302±19,598	4.3	121,021±23,461	5.4	62,571±30,199	3.0	112,282±36,388	4.2	112,163±22,420	2.1
0.30%	325,485±46,981	12.0	241,465±73,709	9.2	354,678±27,371	10.0	365,768±51,573	8.5	166,224±43,333	14.0	210,636±46,213	15.1	296,024±33,270	13.2	259,203±105,308	12.6	292,230±5423	10.9	251,172±40,569	4.7
B: Hexyl cinnamic aldehyde																				
AOO	24,583±5761	–	41,189±17,452	–	35,652±12,253	–	43,007±8931	–	19,146±6582	–	16,375±3953	–	29,925±6142	–	12,207±4127	–	29,602±8049	–	29,077±2876	–
5%	33,196±6535	1.4	56,291±5484	1.4	48,383±14,959	1.4	64,212±6709	1.5	23,417±6260	1.2	27,369±8594	1.7	46,148±14,005	1.5	16,616±4630	1.4	25,602±11,242	0.9	40,685±14,674	1.4
10%	73,884±14,255	3.0	109,204±15,298	2.7	82,040±12,032	2.3	138,873±51,932	3.2	35,432±14,357	1.9	38,327±9530	2.3	126,755±35,639	4.2	50,829±8197	4.2	65,640±27,871	2.2	79,321±10,548	2.7
25%	142,130±29,633	5.8	198,520±40,800	4.8	158,304±26,958	4.4	219,687±29,834	5.1	76,029±5733	4.0	90,067±27,828	5.5	212,285±50,835	7.1	124,803±34,287	10.2	114,791±13,669	3.9	101,984±21,546	3.5

C: 3-Aminophenol

Vehicle/concentration	1		3		8	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	27,188±10,027	–	24,047±3932	–	20,576±5546	–
1%	47,591±2668	1.8	33,875±4945	1.4	25,167±4299	1.2
3%	63,021±9400	2.3	42,352±11,487	1.8	40,921±10,896	2.0
10%	76,927±15,323	2.8	41,759±8243	1.7	49,037±8244	2.4

D: Glutaraldehyde

Vehicle/concentration	1		2		5	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
ACE	17,947±4929	–	38,044±13,217	–	16,439±6488	–
0.05%	25,594±9403	1.4	28,096±9168	0.7	17,024±5163	1.0
0.15%	72,748±20,584	4.1	48,980±8745	1.3	40,319±17,078	2.5
0.50%	89,767±21,798	5.0	129,110±31,985	3.4	42,237±6048	2.6

E: Cobalt chloride

Vehicle/concentration	4		6		8	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	100,396±24,632	–	4184±2395	–	19,803±4451	–
0.30%	–±–	–	44,002±30,922	10.5	87,562±13,336	4.4
1.00%	203,895±24,479	2.0	44,465±23,293	10.6	131,004±34,534	6.6
3.00%	267,172±52,088	2.7	85,978±24,933	20.6	159,808±13,473	8.1

F: Isoeugenol

Vehicle/concentration	4		5		9	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	42,866±9956	–	11,899±7366	–	26,842±9515	–
1%	125,838±22,236	2.9	22,896±7449	1.9	69,256±20,292	2.6
3%	175,277±10,289	4.1	23,619±8830	2.0	86,598±20,489	3.2
10%	262,118±34,406	6.1	117,098±5209	9.8	190,392±38,486	7.1

G: Formaldehyde

Vehicle/concentration	1		2		5	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
ACE	17,947±4929	–	38,044±13,217	–	16,439±6488	–
0.5%	52,214±10,965	2.9	64,467±11,056	1.7	19,510±5239	1.2
1.5%	51,405±13,007	2.9	115,143±20,638	3.0	30,959±12,804	1.9
5.0%	86,934±33,682	4.8	120,966±21,688	3.2	44,219±7822	2.7

H: Dimethyl isophthalate

Vehicle/concentration	1		3		7	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	27,188±10,027	–	35,610±7212	–	22,466±3515	–
5%	36,534±10,199	1.3	35,710±8126	1.0	28,306±4047	1.3
10%	31,200±10,875	1.1	34,357±8364	1.0	25,555±3074	1.1
25%	30,030±10,456	1.1	23,900±3733	0.7	23,583±3751	1.0

I: Isopranol

Vehicle/concentration	1		2		3		4		5		6		7		8		9		10	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	24,583±5761	–	41,189±17,452	–	35,652±12,253	–	43,007±8931	–	19,146±6582	–	16,375±3953	–	29,925±6142	–	12,207±4127	–	29,602±8049	–	29,077±2876	–
10%	37,756±12,448	1.5	37,286±9163	0.9	36,155±7444	1.0	67,307±12,946	1.6	10,106±3170	0.5	32,233±26,281	2.0	43,446±17,986	1.5	14,797±2984	1.2	18,791±7645	0.6	26,480±4594	0.9
25%	27,101±2623	1.1	35,024±4878	0.9	23,465±7953	0.7	38,859±7172	0.9	14,531±1549	0.8	14,762±5342	0.9	27,285±10,469	0.9	12,387±3421	1.0	20,627±6175	0.7	30,676±5707	1.1
50%	28,723±3313	1.2	33,259±7651	0.8	30,822±4465	0.9	34,382±6421	0.8	13,581±3696	0.7	21,360±4957	1.3	24,776±3613	0.8	13,551±1001	1.1	15,039±3697	0.5	36,216±4960	1.2

J: Nickel sulfate

Vehicle/concentration	4		6		8	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	100,396±24,632	–	4184±2395	–	19,803±4451	–
1%	116,266±22,468	1.2	21,990±7141	5.3	69,077±14,602	3.5
3%	153,074±35,051	1.5	27,966±6162	6.7	60,881±7880	3.1
10%	103,595±20,343	1.0	49,303±14,901	11.8	50,568±9846	2.6

K: Abietic acid

Vehicle/concentration	2		6		7	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	26,159±2157	–	13,910±3921	–	21,546±13,493	–
5%	55,039±8805	2.1	25,277±9139	1.8	40,328±8389	1.9
10%	91,706±17,069	3.5	57,615±12,621	4.1	85,821±24,030	4.0
25%	121,351±36,474	4.6	110,697±29,265	8.0	81,818±24,819	3.8

L: Methyl salicylate

Vehicle/concentration	3		7		10	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	24,047±3932	–	21,546±13,493	–	53,350±14,893	–
5%	25,764±7330	1.1	23,459±7751	1.1	33,663±5192	0.6
10%	26,361±6381	1.1	38,158±6803	1.8	41,698±7559	0.8
25%	37,359±10,622	1.6	29,881±11,569	1.4	44,426±13,600	0.8

(continued on next page)

Table 3(a) (*continued*)

Positive control (hexyl cinnamic aldehyde)

Vehicle/ concentration	1		2		3		4		5		6		7		8		9		10	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	23,639±5906	–	30,284±11,576	–	25,429±5894	–	44,371±9224	–	15,183±5554	–	10,447±4413	–	25,112±8035	–	18,428±4503	–	26,327±5484	–	22,309±6393	–
25%	147,032±30,059	6.2	153,995±35,670	5.1	144,091±18,550	5.7	243,877±42,495	5.5	72,877±19,820	4.8	84,748±16,459	8.1	136,327±26,932	5.4	101,382±22,894	5.5	140,388±23,895	5.3	113,209±18,835	5.1

ACE, acetone; AOO, acetone–olive oil; DMSO, dimethylsulfoxide.
Number of animals: 4 for all the tested chemicals, 8 for the positive controls of laboratories 9 and 10, and 12 for the positive controls of laboratories 1–8.

the same vehicle control group were employed. These 4 chemicals were randomly allocated by a biostatistician.

In order to avoid predicting the severity of the effects of each chemical, all the chemical names were coded into alphabetic characters, and they were labeled as low, medium, and high in terms of the concentration that enabled blinded distribution for both the studies. 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.

2.5. Development of LLNA-DA

The original LLNA measures the proliferation of draining lymph node cells (LNCs) via the incorporation of [^{3}H]-methyl thymidine into DNA and β scintillation counting. Although this approach to measure the activity of LNC is well established through many studies on the original LLNA, alternative approaches that do not require the use of radioisotopes are expected to be beneficial.

ATP is the main energy source for a majority of cellular functions, and it is an essential molecule for living cells. ATP activity is known to indicate the number of living cells. Therefore, measurement of the ATP content in the lymph node by a luciferin–luciferase assay is considered to be one of the surrogates of altered lymph node cellularity. The measurement of the ATP content of the lymph node involves determination of the cell number at the end of cell proliferation, while the measurement of [^{3}H]-methyl thymidine incorporation involves determination of the endpoint of cell proliferation. One of the benefits of measuring the ATP content is that it allows the use of commercially available reagent kits; in this method, the ATP content is expressed in terms of the chemiluminescence (relative light units, RLU) induced by the luciferin–luciferase reaction.

Yamashita, Idehara, Fukuda, Yamagishi, and Kawada (2005) used 3 chemicals to study the approach involving the measurement of the ATP content. They found that when the dosing schedule of the original LLNA was followed, the ATP measurement approach as well as the flow cytometric analysis of LNCs (Hatao, Hariya, Katsumura, & Kato, 1995) or the assessment of 5-bromo-2′-deoxyuridine (BrdU) incorporation into LNCs (Takeyoshi, Yamasaki, Yakabe, Takatsuki, & Kimber, 2001) tended to show lower stimulation indices (SIs) than the original LLNA. Hence, in order to increase lymph node proliferation, Yamashita et al. proposed pretreatment with 1% sodium lauryl sulfate (SLS) prior to the application of the test chemicals and an additional treatment with the tested chemical. Through their studies, these authors successfully increased the sensitivity of the ATP measurement approach, and the SI value of 3 obtained with this approach was considered to be comparable to that of the original LLNA. Additionally, these authors conducted 6 independent experiments using eugenol to determine the intralaboratory variation in the SI values of the ATP measurement approach. The mean and coefficient of variance of the SI values were 4.0% and 17.3%, respectively.

Daicel Chemical Industries Ltd. refined the ATP measurement approach, which was designated LLNA-DA. In addition to the original LLNA procedure, this ATP content measurement assay includes pretreatment with 1% SLS solution along with its application of the test chemicals on the seventh day; this strategy was expected to yield similar SI values, i.e., approximately 3, to those of the original LLNA. Therefore, this additional step enabled the use of the same cut-off point as that of the original LLNA. By the time the first validation study was conducted, Daicel Chemical Industries Ltd. had obtained some results for LLNA-DA by using the abovementioned cut-off point, in which the correlation coefficient of the EC3 value for LLNA and LLNA-DA for 10 chemicals was 0.90,

Table 3(b)
Mean and SD for the ATP content and SI values obtained in all the laboratories in the second study

B: Hexyl cinnamic aldehyde

Vehicle/ concentration	11		12		13		14		15		16		17	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	21,328±8537	–	27,436±7629	–	24,739±6350	–	24,348±8236	–	31,189±10,511	–	28,421±8943	–	23,888±10,275	–
5%	32,306±7470	1.5	45,178±8970	1.6	35,059±13,111	1.4	50,408±15,075	2.1	46,853±7275	1.5	65,209±12,332	2.3	31,668±6045	1.3
10%	70,689±7059	3.3	94,494±20,913	3.4	110,638±34,223	4.5	88,935±49,202	3.7	78,471±11,510	2.5	146,720±30,935	5.2	110,331±13,800	4.6
25%	95,348±32,502	4.5	156,615±19,035	5.7	133,833±22,340	5.4	185,142±43,204	7.6	122,146±25,678	3.9	239,220±35,785	8.4	154,106±28,583	6.5

E: Cobalt chloride

Vehicle/concentration	11		13		14		17	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	82,093±26,296	–	81,326±13,350	–	41,770±12,971	–	50,815±5671	–
1%	122,193±21,742	1.5	133,890±34,318	1.6	97,101±15,349	2.3	148,776±68,574	2.9
3%	141,919±33,024	1.7	199,335±5756	2.5	171,272±19,452	4.1	216,116±18,966	4.3
5%	165,350±10,204	2.0	206,394±16,349	2.5	177,705±46,577	4.3	256,978±54,531	5.1

J: Nickel sulfate

Vehicle/concentration	11		12		14		16	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	82,093±26,296	–	83,046±6308	–	41,770±12,971	–	76,153±28,228	–
1%	53,652±8085	0.7	82,896±14,003	1.0	77,804±25,666	1.9	90,029±11,264	1.2
3%	65,034±25,414	0.8	103,345±24,614	1.2	65,200±11,620	1.6	118,932±13,811	1.6
10%	60,451±17,784	0.7	80,596±21,515	1.0	88,990±14,982	2.1	88,482±19,237	1.2

M: Lactic acid

Vehicle/concentration	11		13		15		16	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	65,060±9211	–	81,326±13,350	–	49,353±21,291	–	76,153±28,228	–
5%	60,576±20,296	0.9	80,639±18,883	1.0	45,730±8622	0.9	69,247±15,579	0.9
10%	49,033±11,761	0.8	55,369±7627	0.7	47,928±15,171	1.0	60,621±11,273	0.8
25%	52,131±16,088	0.8	60,124±13,945	0.7	35,259±2939	0.7	69,108±14,746	0.9

N: Potassium dichromate

Vehicle/concentration	11		12		15		16	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
DMSO	65,060±9211	–	83,046±6308	–	49.353±21,291	–	50,815±5671	–
01%	123,936±17,967	19	157,464±29,682	19	131,244±35,222	27	165,248±46,056	33
0.3%	145,833±41,893	2.2	217,061±37,807	2.6	191,819±51,627	3.9	257,138±29,816	5.1
1.0%	311,009±24,188	4.8	338,610±33,485	4.1	296,431±75,377	6.0	323,834±60,878	6.4

Positive control (hexyl cinnamic aldehyde)

Vehicle/ concentration	11		12		13		14		15		16		17	
	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI	Mean±SD	SI
AOO	25,807±8795	–	30,147±6951	–	24,943±6509	–	27,245±7022	–	33,713±7937	–	37,383±5294	–	17,417.3±7195	–
25%	102,118±22,127	4.0	142,679±50,388	4.7	136,950±22,057	5.5	184,010±31,146	6.8	143,322±31,990	4.3	268,199±47,663	7.2	138,799±25,305	8.0

ACE, acetone; AOO, acetone–olive oil; DMSO, dimethylsulfoxide.
Number of animals: 4 for all the tested chemicals, 8 for the positive controls of laboratories 12–17, and 12 for the positive control of laboratory 11.

and the accuracy of LLNA-DA against LLNA for 18 chemicals was 89% (16/18) (in-house data).

The ATP content value is influenced by time, that is, it decreases over time. This is not emerge in the original LLNA since it involves the measurement of [^{3}H]-methyl thymidine incorporation. Daicel Chemical Industry Ltd. investigated it and found that the ATP content value is not influenced by a 10- to 20-min delay, while this value would be reduced to approximately 50% of its original value with a 2-h delay. Therefore, Daicel Chemical Industry Ltd. recommends that when LLNA-DA is conducted, all the procedural steps from lymph node excision to the determination of the ATP content be performed rapidly and without delay.

Very recently, Idehara et al. (in press) reported the details of the intralaboratory study on LLNA-DA.

2.6. Standard protocol of LLNA-DA for the studies

The standard protocol for the assay was prepared prior to the preliminary test and determined according to the time of commencement of the study. Three doses were prepared for each of the test chemicals.

The groups of female CBA/JNCrlj mice (n=4; Charles River Japan Inc., Kanagawa) were treated with the topical application of 25 μL of 1 of the 3 doses of the test chemicals or the vehicle control exclusively on the dorsum of both ears. Following pretreatment with 1% SLS for 1 h, daily treatments with the chemicals were performed for the first 3 days and, subsequently, on day 7. On day 8, the treated mice were sacrificed, and the draining auricular lymph nodes were excised. After recording the lymph node weight (LNW), the LNCs were ground

between 2 slide glasses and subsequently suspended in 1 mL of phosphate-buffered saline (PBS) with a cell scraper. The LNC suspension was mixed and diluted to 1% with PBS. The ATP content was determined using a commercially available kit (Kikkoman Co., Tokyo). ATP was extracted from 0.1 mL of the diluted LNC suspension for 20 s, following which 0.1 mL of a reagent containing luciferase was added and the bioluminescence (RLU) in 10 s was measured with a luminometer (Lumitester C-100; Kikkoman Co., Tokyo). A point to note is that after the death of the animal, the ATP content of the lymph node decreases over time. It is therefore desirable that the series of procedures from lymph node excision to the determination of the ATP content must be performed rapidly and without delay.

2.7. Database

A biostatistician created a database containing the LNW and ATP content data obtained for each mouse in all the experimental laboratories. For comparison, data from studies on the original LLNA were collected and included in the database.

2.8. Statistical methods

For each experimental group, the SI was defined as the increase in the ATP content in the chemical-treated group relative to that in the vehicle control group. An SI of 3 was defined as the cut-off value for the skin sensitization potential. In order to demonstrate the variability within the SI values, the confidence interval of the SI values was calculated (Omori & 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 interlaboratory variations; this is similar to the meta-analysis technique used in clinical studies (Normand, 1999). Using the abovementioned random effect model, we estimated the weighted average as an overall estimate of the SI value recorded for each chemical dose. The EC3 is defined as the estimated concentration that yields an SI value of 3. The EC3 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-DA results with the LLNA or GPMT/BT results (OECD, 2005). These measures were not calculated in the second study because of a shortage of chemicals.

3. Results

3.1. Chemical selection

Tables 1(a) and 1(b) show the selected chemicals, the results of LLNA and GPMT/BT as references, and the results obtained for the chemicals allocated for the LLNA-DA experiments of both the studies.

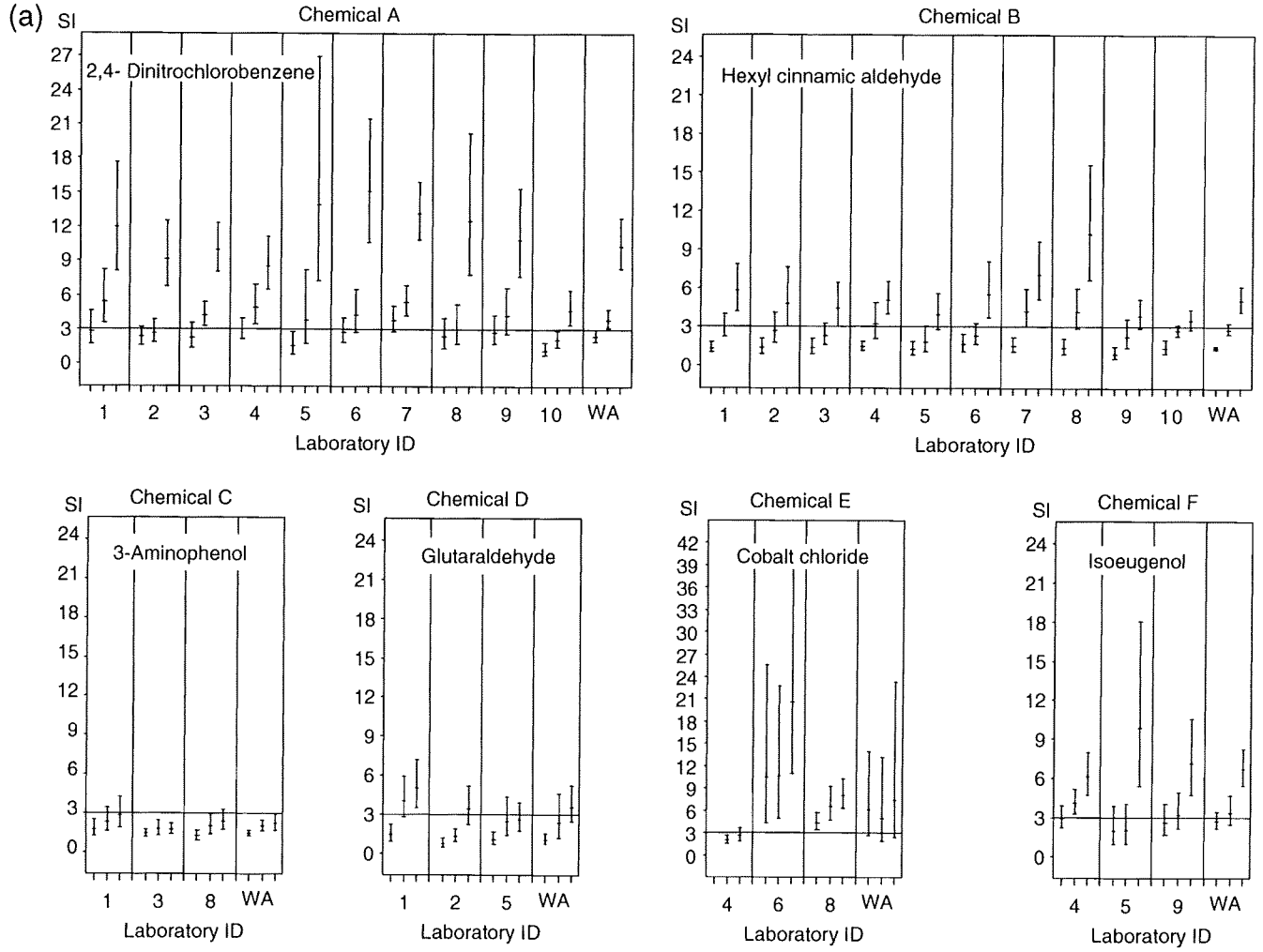

Fig. 1. (a). Dose–response relationships of the SI values with 95% confidence intervals for each chemical analyzed in all the laboratories. "WA" indicates the weighted average of the SI values obtained by meta-analysis using the random effect model in the first study. (b). Dose–response relationships of the SI values with 95% confidence intervals for each chemical analyzed in all the laboratories. "WA" indicates the weighted average of the SI values obtained by meta-analysis using the random effect model in the second study.

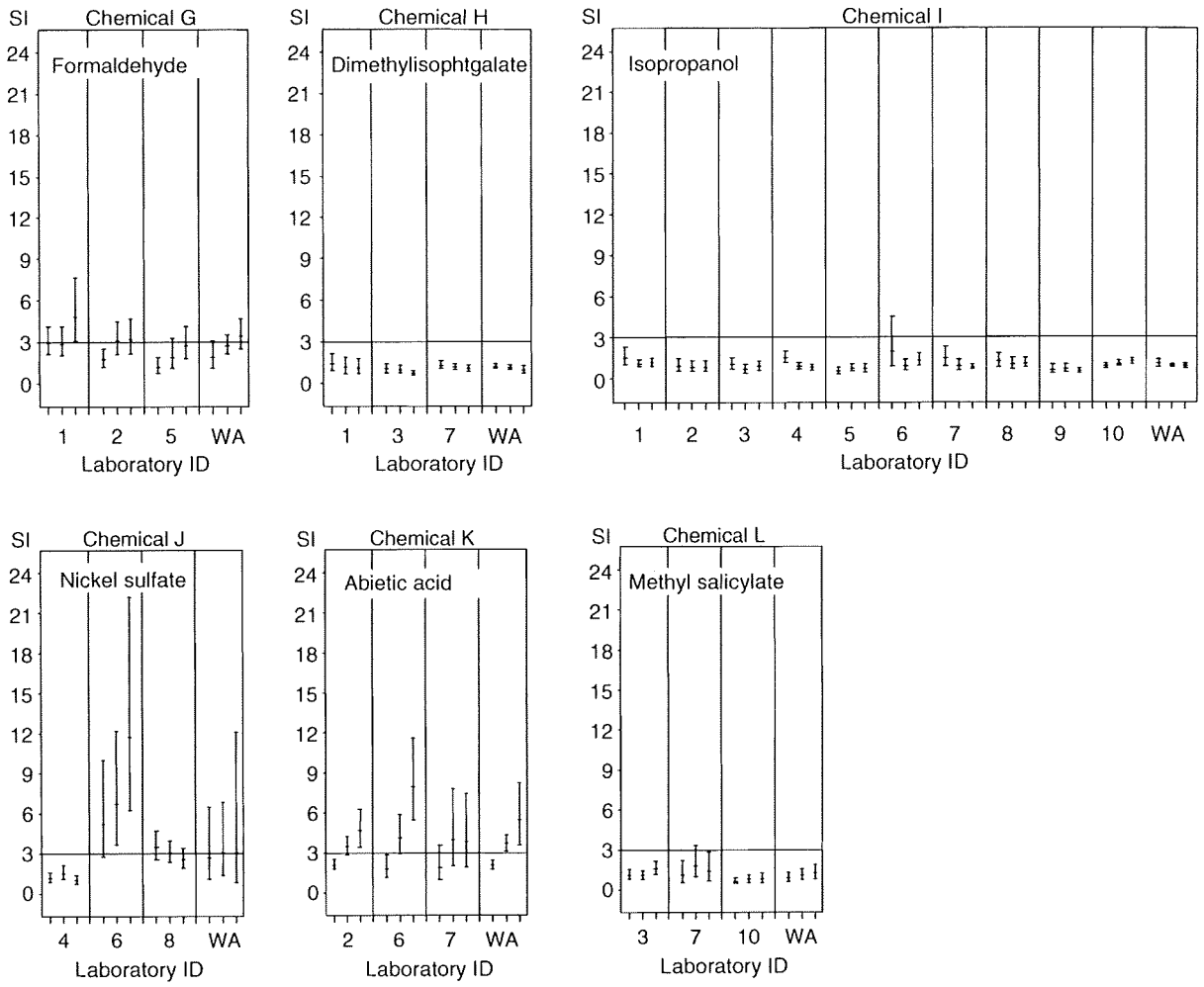

Fig. 1 (*continued*).

The GPMT/BT results for chemical D (glutaraldehyde) are not listed in Table 1(a) because the data were not available at the time the list was prepared.

The chemical selectors initially set the dose concentrations of chemical E (cobalt chloride) at 1%, 3%, and 10%. However, during the first round of the experiments in a laboratory in the first study, 2 of the 4 mice treated with the 10% dose concentration died, while the other 2 exhibited signs of hypokinesia. Since only the laboratory had conducted the experiment using this chemical concentration at the time, the chemical selectors decided to alter the dose concentrations. Then, the dose concentrations of chemical E were subsequently set at 0.3%, 1%, and 3% in a blinded manner for the remaining 2 laboratories in the first study. However, after several considerations, the chemical selectors adopted different doses in the second study, i.e., 1%, 3%, and 5%.

3.2. Body weights

Tables 2(a) and 2(b) summarize the body weight statistics observed on days 1 and 8 in each laboratory, respectively. No substantial interlaboratory variations were observed with regard to the body weights.

3.3. ATP content and SI values

The ATP content and SI values recorded by the experimental laboratories for each of the test chemicals are summarized in Tables 3(a) and 3(b), and the dose–response relationships for the SI values are indicated in Fig. 1(a) and (b).

The results of the first study are shown in Table 3(a) and Fig. 1(a). For chemicals A (2,4-dinitrochlorobenzene), B (hexyl cinnamic aldehyde), F (isoeugenol), and K (abietic acid), dose–response relationships of the SI values were clearly evident in each laboratory, and the SI values for all the high-dose groups were greater than 3. The dose–response relationships for chemicals H (dimethyl isophthalate), I (isopropanol), and L (methyl salicylate) were unclear, and the laboratories that assessed these chemicals reported negative findings. The SI values obtained for chemical C (3-aminophenol) in all 3 laboratories were lower than 3, and the values obtained in laboratories 1 and 3 were approximately 3 for the high-dose group. Further, dose–response relationships of the SI values were observed for chemicals D (glutaraldehyde) and G (formaldehyde), whose SI values were also approximately 3 for the high-dose groups. The SI values were greater than 3 for the high-dose groups in laboratories 1 and 2 but not in laboratory 5. The SI values for chemicals E (cobalt chloride) and J (nickel sulfate) were inconsistent across laboratories; further, an inconsistency was observed in the ATP content values in the vehicle control group for these chemicals. In the case of chemical E, the dose–response relationship of the weighted average of the SI values yielded a v-shaped curve; therefore, it may be considered that the observed dose–response relationships based on the weighted average values for chemical E were inappropriate.

Table 3(b) and Fig. 1(b) describe the results of the second study. For chemicals B (hexyl cinnamic aldehyde) and N (potassium dichromate),

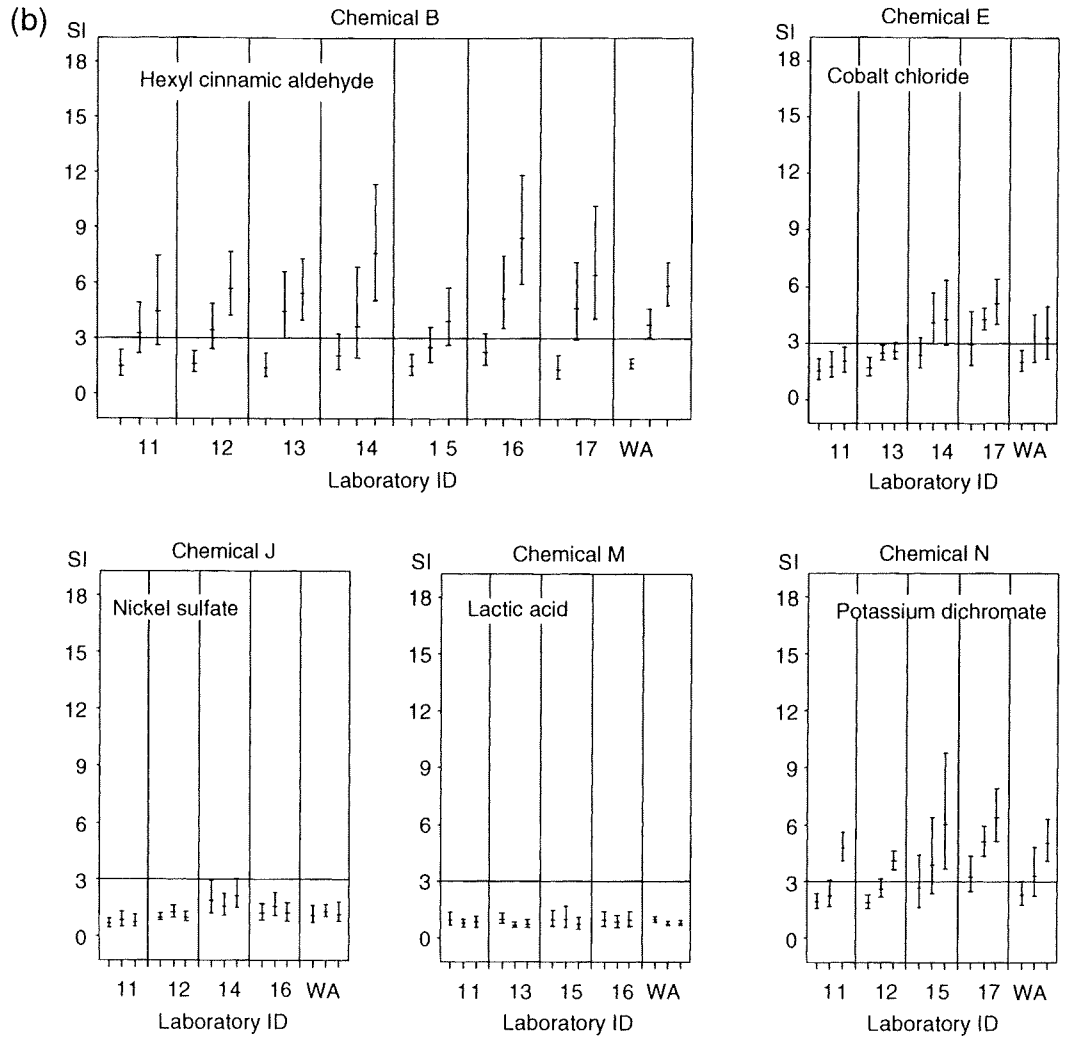

Fig. 1 (*continued*).

the dose–response relationships of the SI values were evident in each laboratory, and all the SI values of the high-dose groups were greater than 3. The SI values for chemicals J (nickel sulfate) and M (lactic acid) were lower than 3, and these chemicals tested negative in all the laboratories. The SI value for chemical E (cobalt chloride), which was inconsistent in the first study, was also inconsistent between different laboratories in the study. However, as opposed to the results of the first study, the dose–response relationships and ATP contents were considerably similar between laboratories.

3.4. ATP content and LNW

Fig. 2(a) and (b) shows the scatter plots of ATP content according to LNW for all the chemicals. Since the ATP content decreases with time, it is important for the scatter plot to demonstrate a linear relationship between the ATP content and LNW. This linear relationship can be used as a rough indicator of whether the experiments conformed to the protocol for measuring the ATP content. Since all the scatter plots demonstrated linearity, it can be concluded that all the experiments adhered to the protocol.

3.5. Assay sensitivity

We defined assay sensitivity as the ability to accurately detect the positive control chemical. Since a positive control was included in each experiment, we investigated whether the SI value assigned to the positive control group was greater than 3 in the experiments. Fig. 3(a) and (b) shows the SI values obtained for all the positive control groups with 95% confidence intervals. All the experiments in these studies were assay sensitive because all the SI values were greater than 3.

3.6. Intralaboratory variability

Although limited, the results obtained for the positive control groups allowed us to evaluate the intralaboratory variability of the assay. Fig. 3(a) and (b) also shows the variability of the SI values obtained for the positive control groups in each laboratory in both the studies. No large intralaboratory variation was observed in any of the laboratories.

3.7. Interlaboratory variability

The data shown in Fig. 1(a) and (b) were used to measure the interlaboratory variability in the SI values for all the chemical doses. Tables 4(a) and 4(b) show the weighted average of the SI values with 95% confidence intervals and a summary index of the interlaboratory variability, i.e., τ^2.

In the first study, all the doses of chemicals E (cobalt chloride) and J (nickel sulfate) and the intermediate dose of chemical D (glutaraldehyde) exhibited relatively large interlaboratory variations. On the

other hand, in the second study, no large interlaboratory variation was observed in any of the laboratories.

Tables 5(a) and 5(b) show the results of the judgments based on the cut-off value of 3 for the SI values obtained for all the chemicals in all the laboratories. In the first study, 4 chemicals, namely, D (glutaraldehyde), E (cobalt chloride), G (formaldehyde), and J (nickel sulfate), showed inconsistent results among the laboratories. For chemicals D (glutaraldehyde) and G (formaldehyde), the SI values for the high doses were approximately 3 among all 3 laboratories; thus, the variation was small. On the other hand, the values for chemicals E (cobalt chloride) and J (nickel sulfate) were inconsistent among the laboratories (Fig. 1(a)). In the second study, consistent results were observed for the 4 chemicals. Although an inconsistency was observed for chemical E (cobalt chloride), the dose-

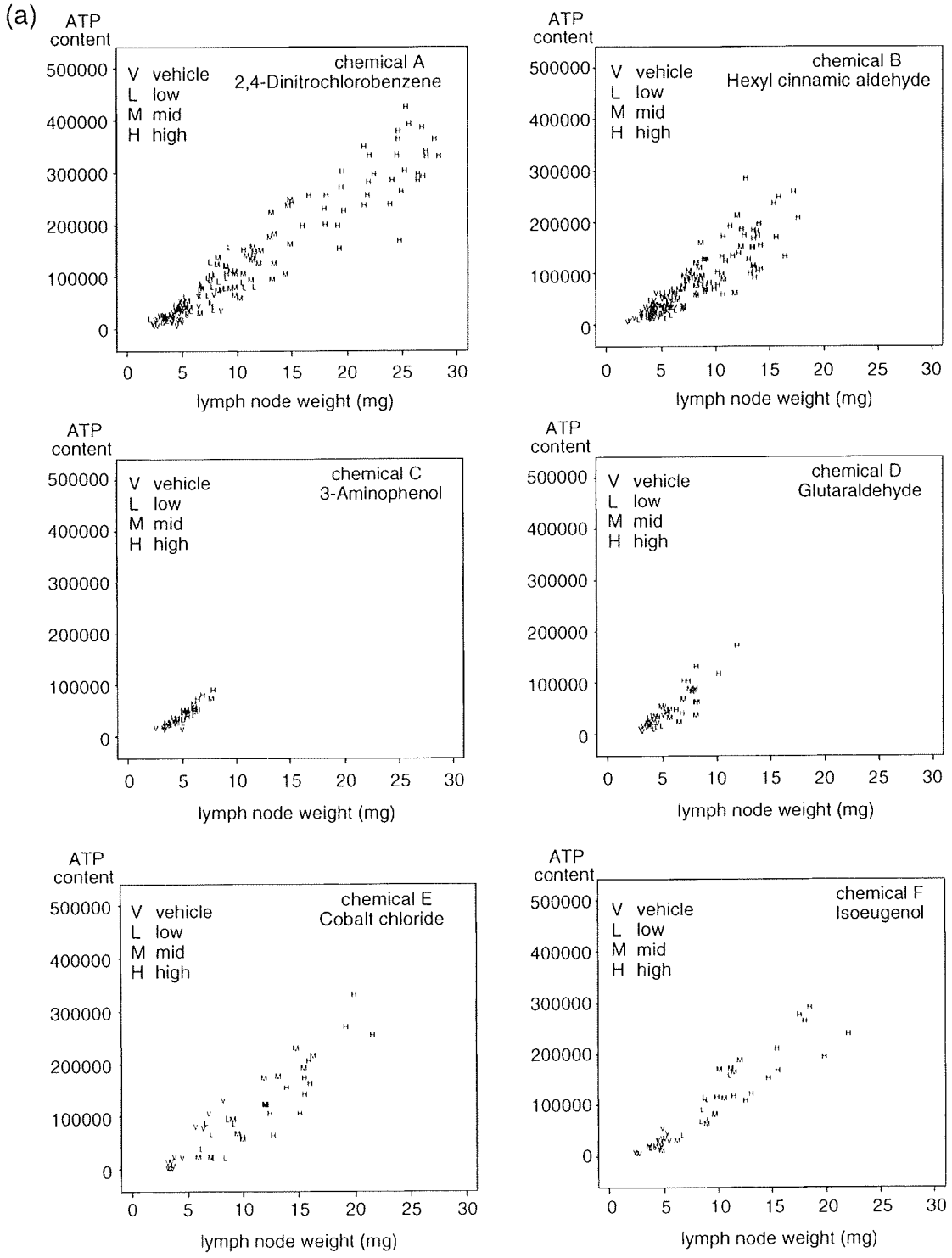

Fig. 2. (a). Scatter plots indicating the ATP content with the LNW (mg) recorded for the vehicle (V), low-dose (L), middle-dose (M), and high-dose groups (H) for each chemical in the first study. (b). Scatter plots indicating the ATP content with the LNW (mg) recorded for the vehicle (V), low-dose (L), middle-dose (M), and high-dose groups (H) for each chemical in the second study.