

Annex V

Accuracy Analyses Using Additional Approaches for Combining Multiple Test Results

This page intentionally left blank

1.0 Accuracy Analyses Using Alternative Decision Criteria and Alternate Methods for Combining Data for Substances Tested Multiple Times

This annex shows performance analyses for the murine local lymph node assay (LLNA) modified by Daicel Chemical Industries, Ltd., based on ATP content (referred to hereafter as the “LLNA: DA”) for alternative decision criteria when using two different approaches for combining test results for the 14 substances with multiple LLNA: DA tests.

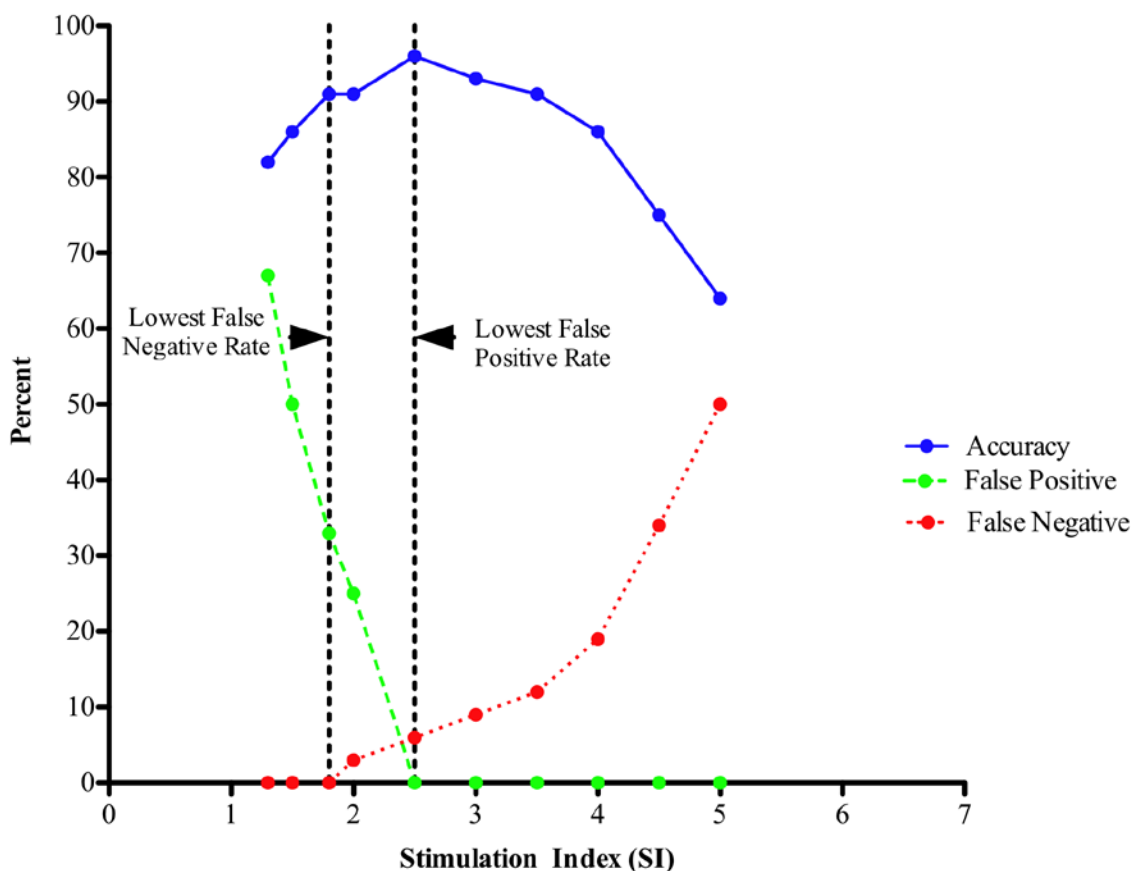
1. The positive/negative outcome for each substance for each criterion was determined by the outcome of the test with the highest maximum stimulation index (SI) of the multiple tests.
2. The positive/negative outcome for each substance for each criterion was determined by the outcome of the test with the lowest maximum SI of the multiple tests.

Section 6.0 of this background review document provides the results for the analysis when the most prevalent outcome was used to represent the result for each substance tested multiple times (for each criterion).

1.1 Results of LLNA: DA Accuracy Analysis Using Alternative Decision Criteria and Highest Maximum SI for the Outcome of Multiple Tests

When combining multiple test results for a single substance by using the outcome of the test with the highest maximum SI to identify potential sensitizers, the decision criterion of $SI \geq 3.0$ (used by the LLNA: DA validation study team) yielded an accuracy of 93% (41/44), a sensitivity of 91% (29/32), a specificity of 100% (12/12), a false positive rate of 0% (0/12), and a false negative rate of 9% (3/32) (**Table C-V-1**). The decision criteria using higher SI values, $SI \geq 3.5$ to $SI \geq 5.0$, decreased performance except for the specificity and the false positive rate, which remained at 100% (12/12) and 0% (0/12), respectively (**Figure C-V-1** and **Table C-V-1**). The lower SI criterion, $SI \geq 1.8$, decreased accuracy to 91% (40/44) but increased sensitivity to 100% (32/32), while the specificity and the false positive rate decreased to 67% (8/12) and 33% (4/12), respectively. Further, the false negative rate decreased to 0% (0/32) at $SI \geq 1.8$. The use of analysis of variance (ANOVA) and summary statistics (i.e., mean ATP measurement of treated groups $\geq 95\%$ confidence interval [CI] of the control group, or ≥ 2 or ≥ 3 standard deviation [SD] from the control group mean), yielded accuracy values of 75% to 84%, with sensitivity values of 88% to 100%, and false negative rates of 0 to 13%. The specificity for these criteria ranged from 8% to 58% and the false positive rates were 42% to 92%. As summarized above, the best overall performance of these alternative decision criteria (based on the highest SI cutoff that yielded no false positives) was achieved using an $SI \geq 1.8$ and using the highest maximum SI for substances with more than one test. Using a cutoff at $SI \geq 1.8$, however, misclassified four nonsensitizers in the traditional LLNA (including isopropanol based on its highest maximum SI of 1.97).

Figure C-V-1 Performance of the LLNA: DA for 44 Substances Compared to the Traditional LLNA in Predicting Skin Sensitization Potential Using the Highest Maximum SI for Substances with Multiple Tests



As compared to traditional LLNA results, the lines show the change in performance characteristics for the LLNA: DA with the SI cutoff used to identify sensitizers. This analysis used LLNA: DA and traditional LLNA results for 44 substances (32 traditional LLNA sensitizers and 12 traditional LLNA nonsensitizers). For the 14 substances with multiple test results, the result for each substance was based on the test with the highest maximum SI value. The solid line shows accuracy, the dashed line shows the false positive rate, and the dotted line shows the false negative rate.

Abbreviations: LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd. based on ATP content; SI = stimulation index.

Table C-V-1 Performance of the LLNA: DA for 44 Substances Compared to the Traditional LLNA in Predicting Skin Sensitization Potential Using Alternative Decision Criteria Based on the Highest Maximum SI for Substances with Multiple Tests

Alternate Criterion	N ¹	Accuracy % (No. ²)	Sensitivity % (No. ²)	Specificity % (No. ²)	False Positive Rate % (No. ²)	False Negative Rate % (No. ²)	Positive Predictivity % (No. ²)	Negative Predictivity % (No. ²)
Statistics ³	44	84 (37/44)	94 (30/32)	58 (7/12)	42 (5/12)	6 (2/32)	86 (30/35)	78 (7/9)
≥95% CI ⁴	44	75 (33/44)	100 (32/32)	8 (1/12)	92 (11/12)	0 (0/32)	74 (32/43)	100 (1/1)
≥2 SD ⁵	44	77 (34/44)	91 (29/32)	42 (5/12)	58 (7/12)	9 (3/32)	81 (29/36)	63 (5/8)
≥3 SD ⁶	44	77 (34/44)	88 (28/32)	50 (6/12)	50 (6/12)	13 (4/32)	82 (28/34)	60 (6/10)
SI ≥ 5.0	44	64 (28/44)	50 (16/32)	100 (12/12)	0 (0/12)	50 (16/32)	100 (16/16)	43 (12/28)
SI ≥ 4.5	44	75 (33/44)	66 (21/32)	100 (12/12)	0 (0/12)	34 (11/32)	100 (21/21)	52 (12/23)
SI ≥ 4.0	44	86 (38/44)	81 (26/32)	100 (12/12)	0 (0/12)	19 (6/32)	100 (26/26)	67 (12/18)
SI ≥ 3.5	44	91 (40/44)	88 (28/32)	100 (12/12)	0 (0/12)	13 (4/32)	100 (28/28)	75 (12/16)
<i>SI ≥ 3.0</i>	<i>44</i>	<i>93 (41/44)</i>	<i>91 (29/32)</i>	<i>100 (12/12)</i>	<i>0 (0/12)</i>	<i>9 (3/32)</i>	<i>100 (29/29)</i>	<i>80 (12/15)</i>
<i>SI ≥ 2.5</i>	<i>44</i>	<i>96 (42/44)</i>	<i>94 (30/32)</i>	<i>100 (12/12)</i>	<i>0 (0/12)</i>	<i>6 (2/32)</i>	<i>100 (30/30)</i>	<i>86 (12/14)</i>
SI ≥ 2.0	44	91 (40/44)	97 (31/32)	75 (9/12)	25 (3/12)	3 (1/32)	91 (31/34)	90 (9/10)
SI ≥ 1.8	44	91 (40/44)	100 (32/32)	67 (8/12)	33 (4/12)	0 (0/32)	89 (32/36)	100 (8/8)
SI ≥ 1.5	44	86 (38/44)	100 (32/32)	50 (6/12)	50 (6/12)	0 (0/32)	84 (32/38)	100 (6/6)
SI ≥ 1.3	44	82 (36/44)	100 (32/32)	33 (4/12)	67 (8/12)	0 (0/32)	80 (32/40)	100 (4/4)

Italicized text indicates the decision criterion chosen by the LLNA: DA validation study team; boldface indicates the single decision criterion that had an overall increased performance in predicting skin sensitization potential when compared to the traditional LLNA (i.e., no false negatives).

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP Content; No. = number; SD = standard deviation; SI = stimulation index

¹ N = Number of substances included in this analysis.

² The proportion on which the percentage calculation is based.

³ Analysis of variance for difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at one dose. The ATP data were log-transformed prior to statistical analyses. For analysis of variance, significance at $p < 0.05$ was further tested by Dunnett's test.

⁴ The mean ATP of at least one treatment group was outside the 95% CI for the mean ATP of the vehicle control group.

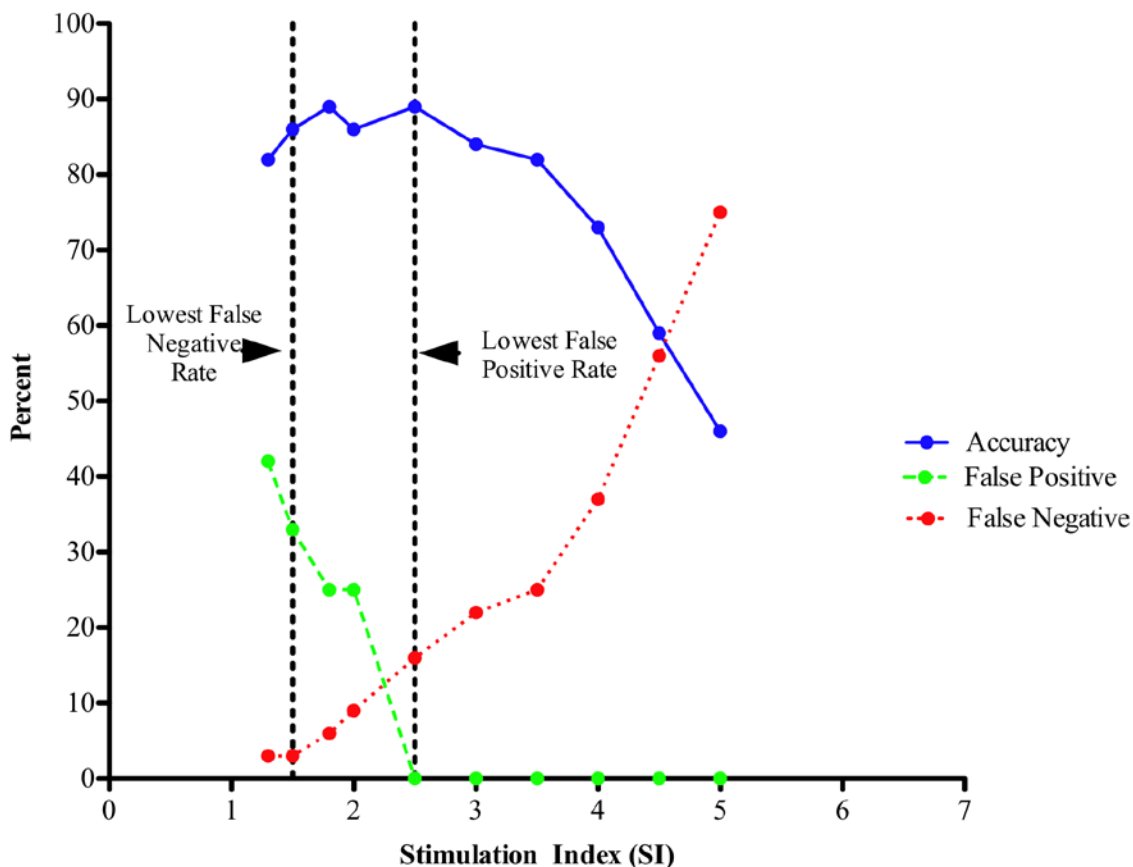
⁵ The mean ATP of at least one treatment group was greater than 2 SD from the mean ATP of the vehicle control group.

⁶ The mean ATP of at least one treatment group was greater than 3 SD from the mean ATP of the vehicle control group.

1.2 Results of LLNA: DA Accuracy Analysis Using Alternative Decision Criteria and Lowest Maximum SI for the Outcome of Multiple Tests

When combining multiple test results for a single substance using the outcome of the test with the lowest maximum SI to identify potential sensitizers, the decision criterion of $SI \geq 3.0$ (used by the LLNA: DA validation study team) yielded an accuracy of 84% (37/44), a sensitivity of 78% (25/32), a specificity of 100% (12/12), a false positive rate of 0% (0/12), and a false negative rate of 22% (7/32) (**Table C-V-2**). The decision criteria using higher SI values, $SI \geq 3.5$ to $SI \geq 5.0$, decreased performance except for the specificity and the false positive rate, which remained at 100% (12/12) and 0% (0/12), respectively (**Figure C-V-2** and **Table C-V-2**). At $SI \geq 5.0$, accuracy decreased to 46% (20/44) and the false negative rate increased to 75% (24/32). Use of a lower SI cutoff at $SI \geq 2.5$ increased accuracy to 89% (39/44) and sensitivity to 84% (27/32), while the specificity and false positive rate remained the same at 100% (12/12) and 0% (0/12), respectively. Further, the false negative rate decreased to 16% (5/32) at $SI \geq 2.5$. At $SI \geq 1.8$, accuracy was unchanged at 89% (39/44) with an increased sensitivity of 94% (30/32) and decreased false negative rate of 6% (2/32), while specificity was 75% (9/12) and the false positive rate was 25% (3/12). At an even lower SI criterion, $SI \geq 1.3$, accuracy was decreased to 86% (38/44) but the sensitivity increased to 97% (31/32), while the specificity was 58% (7/12) and the false positive rate was 42% (5/12). Further, the false negative rate decreased to 3% (1/32) at $SI \geq 1.3$. Use of a statistical test (i.e., ANOVA or *t*-test) and summary statistics (i.e., mean ATP measurements of treated groups $\geq 95\%$ CI of the control group, or ≥ 2 or ≥ 3 SD from the control group mean), yielded accuracy values of 77 to 82%, with sensitivity values of 84 to 97%, and false negative rates of 3 to 16%. Both the specificity and false positive rate for these criteria ranged from 42 to 58%. Of these alternative decision criteria, the best overall performance (i.e., lowest combined false positive and false negative rate) for the approach using the lowest maximum SI for substances with more than one test was achieved using $SI \geq 1.8$, as summarized above.

Figure C-V-2 Performance of the LLNA: DA for 44 Substances Compared to the Traditional LLNA in Predicting Skin Sensitization Potential Using the Lowest Maximum SI for Substances with Multiple Tests



As compared to traditional LLNA results, the lines show the change in performance characteristics for the LLNA: DA with the SI cutoff used to identify sensitizers. This analysis used LLNA: DA and traditional LLNA results for 44 substances (32 traditional LLNA sensitizers and 12 traditional LLNA nonsensitizers). For the 14 substances with multiple test results, the result for each substance was based on the test with the lowest maximum SI value. The solid line shows accuracy, the dashed line shows the false positive rate, and the dotted line shows the false negative rate.

Abbreviations: LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd. based on ATP content; SI = stimulation index.

Table C-V-2 Performance of the LLNA: DA for 44 Substances Compared to the Traditional LLNA in Predicting Skin Sensitization Potential Using Alternative Decision Criteria Based on the Lowest Maximum SI for Substances with Multiple Tests

Alternate Criterion	N ¹	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate		Positive Predictivity		Negative Predictivity	
		%	No. ²	%	No. ²	%	No. ²	%	No. ²	%	No. ²	%	No. ²	%	No. ²
Statistics ³	44	82	36/44	91	29/32	58	7/12	42	5/12	9	3/32	85	29/34	70	7/10
≥95% CI ⁴	44	82	36/44	97	31/32	42	5/12	58	7/12	3	1/32	82	31/38	83	5/6
≥2 SD ⁵	44	77	34/44	88	28/32	50	6/12	50	6/12	13	4/32	82	28/34	60	6/10
≥3 SD ⁶	44	77	34/44	84	27/32	58	7/12	42	5/12	16	5/32	84	27/32	58	7/12
SI ≥ 5.0	44	46	20/44	25	8/32	100	12/12	0	0/12	75	24/32	100	8/8	33	12/36
SI ≥ 4.5	44	59	26/44	44	14/32	100	12/12	0	0/12	56	18/32	100	14/14	40	12/30
SI ≥ 4.0	44	73	32/44	63	20/32	100	12/12	0	0/12	38	12/32	100	20/20	50	12/24
SI ≥ 3.5	44	82	36/44	75	24/32	100	12/12	0	0/12	25	8/32	100	24/24	60	12/20
<i>SI ≥ 3.0</i>	<i>44</i>	<i>84</i>	<i>37/44</i>	<i>78</i>	<i>25/32</i>	<i>100</i>	<i>12/12</i>	<i>0</i>	<i>0/12</i>	<i>22</i>	<i>7/32</i>	<i>100</i>	<i>25/25</i>	<i>63</i>	<i>12/19</i>
<i>SI ≥ 2.5</i>	<i>44</i>	<i>89</i>	<i>39/44</i>	<i>84</i>	<i>27/32</i>	<i>100</i>	<i>12/12</i>	<i>0</i>	<i>0/12</i>	<i>16</i>	<i>5/32</i>	<i>100</i>	<i>27/27</i>	<i>71</i>	<i>12/17</i>
SI ≥ 2.0	44	86	38/44	91	29/32	75	9/12	25	3/12	9	3/32	91	29/32	75	9/12
SI ≥ 1.8	44	89	39/44	94	30/32	75	9/12	25	3/12	6	2/32	91	30/33	82	9/11
SI ≥ 1.5	44	89	39/44	97	31/32	67	8/12	33	4/12	3	1/32	89	31/35	89	8/9
SI ≥ 1.3	44	86	38/44	97	31/32	58	7/12	42	5/12	3	1/32	86	31/36	88	7/8

Italicized text indicates the decision criterion chosen by the LLNA: DA validation study team; boldface indicates the single decision criterion that had an overall increased performance in predicting skin sensitization potential when compared to the traditional LLNA (i.e., lowest combined false positive and false negative rate).

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP Content; No. = number; SD = standard deviation; SI = stimulation index.

¹ N = Number of substances included in this analysis.

² The proportion on which the percentage calculation is based.

³ Analysis of variance for difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at one dose. The ATP data were log-transformed prior to statistical analyses. For analysis of variance, significance at $p < 0.05$ was further tested by Dunnett's test.

⁴ The mean ATP of at least one treatment group was outside the 95% confidence interval for the mean ATP of the vehicle control group.

⁵ The mean ATP of at least one treatment group was greater than 2 SD from the mean ATP of the vehicle control group.

⁶ The mean ATP of at least one treatment group was greater than 3 SD from the mean ATP of the vehicle control group.

.

2.0 Discordant Results for Accuracy Analyses Using Alternative Decision Criteria

This section discusses the discordant results obtained for the analyses using the alternative decision criteria shown in **Tables C-V-1** and **C-V-2**. Discordant results using alternative decision criteria and the highest maximum SI outcome for multiple tests are discussed first (**Section 2.1**), followed by discussion of discordant results using alternative decision criteria and lowest maximum SI outcome for multiple tests (**Section 2.2**). In all cases, discordant results for the alternative decision criteria are discussed using the traditional LLNA as the reference test.

2.1 Discordant Results Using Alternative Decision Criteria and Highest Maximum SI Outcome for Multiple Tests

Table C-V-3 shows how the number and identity of discordant substances changes with the alternative decision criteria when using the test with the highest maximum SI to represent the outcome for substances with multiple tests. Using the decision criterion of $SI \geq 3.0$ to identify sensitizers and the test with the highest maximum SI as the representative result for substances with multiple tests yielded three discordant substances (i.e., 3-aminophenol, 2-mercaptobenzothiazole, and methyl methacrylate), all misclassified as nonsensitizers by the LLNA: DA. Using an SI cutoff lower than three to identify sensitizers, such as $SI \geq 2.0$, yielded four discordant substances: chlorobenzene, hexane, and salicylic acid were misclassified as sensitizers and methyl methacrylate was misclassified as a nonsensitizer. As mentioned in **Section 1.1**, using the decision criterion of $SI \geq 1.8$ to identify sensitizers (based on the test with the highest maximum SI for substances with multiple test results) yielded the highest SI cutoff with no false negatives among the alternative decision criteria evaluated. Yet, when compared to the traditional LLNA, four substances (chlorobenzene, hexane, isopropanol, and salicylic acid) were misclassified as sensitizers by the LLNA: DA. Using a lower SI cutoff of $SI \geq 1.3$ to identify sensitizers, yielded eight discordant substances that were all misclassified as sensitizers (i.e., 1-bromobutane, dimethyl isophthalate, methyl salicylate, and nickel [II] chloride plus the four substances misclassified at $SI \geq 1.8$). Increasing the SI cutoff to values greater than three increased the number of sensitizers that were misclassified as nonsensitizers. At $SI \geq 4.0$, six traditional LLNA sensitizers were misclassified as nonsensitizers by the LLNA: DA while at $SI \geq 5.0$, 16 sensitizers were misclassified as nonsensitizers (**Table C-V-3**).

Use of a statistical test (i.e., ANOVA or *t*-test) or summary statistics (i.e., $\geq 95\%$ CI, ≥ 2 SD, or ≥ 3 SD) tended to misclassify nonsensitizers in the traditional LLNA as sensitizers in the LLNA: DA. Using ANOVA or *t*-test to identify sensitizers misclassified five nonsensitizers (i.e., 1-bromobutane, chlorobenzene, hexane, salicylic acid, and sulfanilamide) as sensitizers and two sensitizers (i.e., 2-mercaptobenzothiazole and methyl methacrylate) as nonsensitizers. Using treatment group ATP measurement with ≥ 2 SD or ≥ 3 SD of the vehicle control mean or a $\geq 95\%$ CI of the vehicle control mean, all misclassified the following six traditional LLNA nonsensitizers as sensitizers: 1-bromobutane, chlorobenzene, hexane, isopropanol, nickel (II) chloride, and propylparaben. The $\geq 95\%$ CI of the vehicle control mean misclassified four additional nonsensitizers (i.e., diethyl phthalate, dimethyl isophthalate, lactic acid, and methyl salicylate) as sensitizers. In addition, ≥ 2 SD or ≥ 3 SD of the vehicle control mean commonly misclassified three sensitizers as nonsensitizers (i.e., ethyl acrylate, methyl methacrylate, and propyl gallate).

Thirteen of the 22 ICCVAM-recommended LLNA performance standards reference substances (ICCVAM 2009) tested in the LLNA: DA were discordant for the analysis of alternative decision criteria using the test with the highest maximum SI to represent substances with multiple tests (**Table C-V-3**) when compared to the traditional LLNA. Six nonsensitizers in the traditional LLNA (i.e., chlorobenzene, isopropanol, lactic acid, methyl salicylate, nickel [II] chloride, and salicylic acid) were misclassified by some criteria in the LLNA: DA as sensitizers, and seven sensitizers in the

traditional LLNA (i.e., citral, ethylene glycol dimethacrylate, imidazolidinyl urea, 2-mercaptobenzothiazole, methyl methacrylate, phenyl benzoate, and sodium lauryl sulfate) were misclassified as nonsensitizers by some criteria when tested in the LLNA: DA.

2.2 Discordant Results Using Alternative Decision Criteria and Lowest Maximum SI Outcome for Multiple Tests

Table C-V-4 shows how the number and identity of discordant substances changes with the alternative decision criteria when using the test with the lowest maximum SI as the representative result for substances with multiple tests. Using an SI cutoff less than three, $SI \geq 2.0$, to identify sensitizers yielded six discordant substances. Three of the six discordant substances (i.e., 3-aminophenol, methyl methacrylate, and nickel [II] sulfate hexahydrate) were misclassified as nonsensitizers by the LLNA: DA compared to the traditional LLNA and the remaining three (i.e., chlorobenzene, hexane, and salicylic acid) were misclassified as sensitizers. As mentioned in **Section 1.2**, using the decision criterion of $SI \geq 1.8$ to identify sensitizers (based on the test with the lowest maximum SI for substances with multiple tests) yielded optimum performance (i.e., lowest combined false positive and false negative rate) for the LLNA: DA when compared to the traditional LLNA. This decision criterion yielded five discordant substances; two were sensitizers in the traditional LLNA but were misclassified as nonsensitizers in the LLNA: DA (i.e., 3-aminophenol and nickel [II] sulfate hexahydrate) and three were nonsensitizers in the traditional LLNA but were misclassified as sensitizers in the LLNA: DA (i.e., chlorobenzene, hexane, and salicylic acid) (**Table C-V-4**).

Using an even lower SI to identify sensitizers, $SI \geq 1.3$, also yielded six discordant substances. Chlorobenzene, hexane, and salicylic acid were still misclassified as sensitizers and nickel (II) sulfate hexahydrate was still misclassified as a nonsensitizer by the LLNA: DA compared to the traditional LLNA. In addition, 1-bromobutane and nickel (II) chloride were also misclassified as sensitizers. Increasing the SI cutoff to values greater than three increased the number of sensitizers that were misclassified as nonsensitizers. At $SI \geq 4.0$, 12 sensitizers were misclassified as nonsensitizers while at $SI \geq 5.0$, 24 sensitizers were misclassified as nonsensitizers. Using the test with the lowest maximum SI as the result for substances with multiple tests caused even potent sensitizers to be misclassified as nonsensitizers at the higher SI cutoffs. For instance, at $SI \geq 5.0$, 2,4-dinitrochlorobenzene and glutaraldehyde were classified as nonsensitizers (**Table C-V-4**).

Use of a statistical test (i.e., ANOVA or *t*-test) or summary statistics (i.e., $\geq 95\%$ CI, ≥ 2 SD, or ≥ 3 SD) more often misclassified traditional LLNA nonsensitizers than sensitizers (**Table C-V-4**). Using ANOVA or *t*-test to identify sensitizers misclassified three sensitizers in the traditional LLNA (i.e., 2-mercaptobenzothiazole, methyl methacrylate, and nickel [II] sulfate hexahydrate) as nonsensitizers in the LLNA: DA. Further, five nonsensitizers in the traditional LLNA (i.e., 1-bromobutane, chlorobenzene, hexane, salicylic acid, and sulfanilamide) were misclassified as sensitizers in the LLNA: DA. Using treatment group ATP measurement $\geq 95\%$ CI, ≥ 2 SD or ≥ 3 SD of vehicle control mean commonly misclassified 1-bromobutane, chlorobenzene, hexane, nickel (II) chloride, and propylparaben as sensitizers and nickel (II) sulfate hexahydrate as a nonsensitizer compared to traditional LLNA results. In addition each summary statistic misclassified from two to four additional substances when compared to traditional LLNA results (see **Table C-V-4**).

Table C-V-3 Discordant Results for the LLNA: DA Using Alternative Decision Criteria Compared to the Traditional LLNA Based on the Highest Maximum SI for Substances with Multiple Tests

Discordant Substance ¹	Alternative Decision Criterion ²													
	Statistics ³	≥95% CI ⁴	≥2 SD ⁵	≥3 SD ⁶	SI ≥ 5.0	SI ≥ 4.5	SI ≥ 4.0	SI ≥ 3.5	SI ≥ 3.0	SI ≥ 2.5	SI ≥ 2.0	SI ≥ 1.8	SI ≥ 1.5	SI ≥ 1.3
3-Aminophenol (3.2%)					-	-	-	-	-					
<i>p</i> -Benzoquinone (0.01%)					-	-	-							
1-Bromobutane (-)	+	+	+	+									+	+
Butyl glycidyl ether (30.9%)				-	-									
Chlorobenzene (-)	+	+	+	+							+	+	+	+
Cinnamic aldehyde (1.9%)					-									
Citral (9.2%)					-	-								
Diethyl maleate (3.6%)					-	-	-							
Diethyl phthalate (-)		+												
Dimethyl isophthalate (-)		+												+
Ethyl acrylate (32.8%)			-	-	-	-								
Ethylene glycol dimethacrylate (28.0%)					-	-								
Hexane (-)	+	+	+	+							+	+	+	+
Imidazolidinyl urea (24.0%)					-									
Isopropanol (-)		+	+	+								+	+	+
Lactic acid (-)		+												
2-Mercaptobenzothiazole (1.7%)	-				-	-	-	-	-	-				

Discordant Substance ¹	Alternative Decision Criterion ²													
	Statistics ³	≥95% CI ⁴	≥2 SD ⁵	≥3 SD ⁶	SI ≥ 5.0	SI ≥ 4.5	SI ≥ 4.0	SI ≥ 3.5	SI ≥ 3.0	SI ≥ 2.5	SI ≥ 2.0	SI ≥ 1.8	SI ≥ 1.5	SI ≥ 1.3
Methyl methacrylate (90.0%)	-		-	-	-	-	-	-	-	-	-			
Methyl salicylate (-)		+											+	+
Nickel (II) chloride (-)		+	+	+										+
Phenyl benzoate (13.6%)					-	-								
Propyl gallate (0.32%)			-	-	-									
Propylparaben (-)		+	+	+										
Resorcinol (6.3%)					-	-								
Salicylic acid (-)	+	+	+								+	+	+	+
Sodium lauryl sulfate (8.1%)					-	-	-	-						
Sulfanilamide (-)	+													
Trimellitic anhydride (4.7%)					-									

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP Content; SD = standard deviation; SI = stimulation index.

¹ Compared to the traditional LLNA; traditional LLNA result in parentheses are “-” for nonsensitizers and EC3 values for sensitizers.

² LLNA: DA outcomes are indicated by “+” for sensitizer results and “-” for nonsensitizer results.

³ Analysis of variance assessed difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at one dose. The ATP data were log-transformed prior to statistical analyses. Significance by analysis of variance at *p* < 0.05 was further tested by Dunnett’s test.

⁴ The mean ATP of at least one treatment group was outside the 95% CI for the mean ATP of the vehicle control group.

⁵ The mean ATP of at least one treatment group was greater than 2 SD from the mean ATP of the vehicle control group.

⁶ The mean ATP of at least one treatment group was greater than 3 SD from the mean ATP of the vehicle control group.

Table C-V-4 Discordant Results for the LLNA: DA Using Alternative Decision Criteria Compared to the Traditional LLNA Based on the Lowest Maximum SI for Substances with Multiple Tests

Discordant Substance ¹	Alternative Decision Criterion ²													
	Statistics ³	≥95% CI ⁴	≥2 SD ⁵	≥3 SD ⁶	SI ≥ 5.0	SI ≥ 4.5	SI ≥ 4.0	SI ≥ 3.5	SI ≥ 3.0	SI ≥ 2.5	SI ≥ 2.0	SI ≥ 1.8	SI ≥ 1.5	SI ≥ 1.3
Abietic Acid (11.9%)					-	-	-							
3-Aminophenol (3.2%)					-	-	-	-	-	-	-	-		
<i>p</i> -Benzoquinone (0.01%)					-	-	-							
1-Bromobutane (-)	+	+	+	+									+	+
Butyl glycidyl ether (30.9%)				-	-									
Chlorobenzene (-)	+	+	+	+							+	+	+	+
Cinnamic aldehyde (1.9%)					-									
Citral (9.2%)					-	-								
Cobalt chloride (0.60%)					-	-	-	-	-	-				
Diethyl phthalate (-)		+												
Dimethyl isophthalate (-)														
Diethyl maleate (3.6%)					-	-	-							
2,4-Dinitrochlorobenzene (0.05%)					-									
Ethyl acrylate (32.8%)			-	-	-	-								
Ethylene glycol dimethacrylate (28.0)					-	-								
Formaldehyde (0.50%)					-	-	-	-	-					
Glutaraldehyde (0.08%)					-	-	-	-	-					

Discordant Substance ¹	Alternative Decision Criterion ²													
	Statistics ³	≥95% CI ⁴	≥2 SD ⁵	≥3 SD ⁶	SI ≥ 5.0	SI ≥ 4.5	SI ≥ 4.0	SI ≥ 3.5	SI ≥ 3.0	SI ≥ 2.5	SI ≥ 2.0	SI ≥ 1.8	SI ≥ 1.5	SI ≥ 1.3
Hexane (-)	+	+	+	+							+	+	+	+
Hexyl cinnamic aldehyde (9.7%)					-	-	-							
Imidazolidinyl urea (24.0%)					-									
2-Mercaptobenzothiazole (1.7%)	-				-	-	-	-	-	-				
Methyl methacrylate (90.0%)	-		-	-	-	-	-	-	-	-	-			
Nickel (II) chloride (-)		+	+	+										+
Nickel (II) sulfate hexahydrate (4.8%)	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Phenyl benzoate (13.6%)					-	-								
Potassium dichromate (0.17%)					-	-								
Propyl gallate (0.32%)			-	-	-									
Propylparaben (-)		+	+	+										
Resorcinol (6.3%)					-	-								
Salicylic acid (-)	+	+	+								+	+	+	+
Sulfanilamide (-)	+													
Sodium lauryl sulfate (8.1%)					-	-	-	-						
Trimellitic anhydride (4.7%)					-									

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: DA = murine local lymph node assay modified by Daicel Chemical Industries, Ltd., based on ATP Content; SD = Standard deviation; SI = Stimulation index.

¹ Compared to the traditional LLNA; traditional LLNA result in parentheses are “-” for nonsensitizers and EC3 values for sensitizers.

² LLNA: DA outcomes are indicated by “+” for sensitizer results and “-” for nonsensitizer results.

³ Analysis of variance for difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at one dose. The ATP data were log-transformed prior to statistical analyses. Significance by analysis of variance at $p < 0.05$ was further tested by Dunnett’s test.

⁴ The mean ATP of at least one treatment group was outside the 95% CI for the mean ATP of the vehicle control group.

⁵ The mean ATP of at least one treatment group was greater than 2 SD from the mean ATP of the vehicle control group.

⁶ The mean ATP of at least one treatment group was greater than 3 SD from the mean ATP of the vehicle control group.

Thirteen of the 22 ICCVAM-recommended LLNA performance standards reference substances (ICCVAM 2009) were discordant for the analysis of alternative decision criteria using the test with the lowest maximum SI as the representative result for substances with multiple tests (**Table C-V-4**). One strong sensitizer in the traditional LLNA, 2,4-dinitrochlorobenzene, was misclassified by $SI \geq 5.0$ as a nonsensitizer in the LLNA: DA. Nine additional sensitizers (i.e., citral, cobalt chloride, ethylene glycol dimethacrylate, hexyl cinnamic aldehyde, imidazolidinyl urea, 2-mercaptobenzothiazole, methyl methacrylate, phenyl benzoate, and sodium lauryl sulfate) were also misclassified as nonsensitizers by some criteria in the LLNA: DA. Three nonsensitizers in the traditional LLNA (i.e., chlorobenzene, nickel [II] chloride, and salicylic acid) were misclassified as sensitizers by some criteria in the LLNA: DA.