

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

Appendix E

LLNA: DA Accuracy Analysis Using Additional Approaches for Combining Multiple Test Results

19

20

21

22

23

24

25

26

27

28

29

[This Page Intentionally Left Blank]

30

30 **1.0 Accuracy Analysis Using Alternative Decision Criteria and**
31 **Alternate Methods for Combining Data for Substances Tested**
32 **Multiple Times**

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

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

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

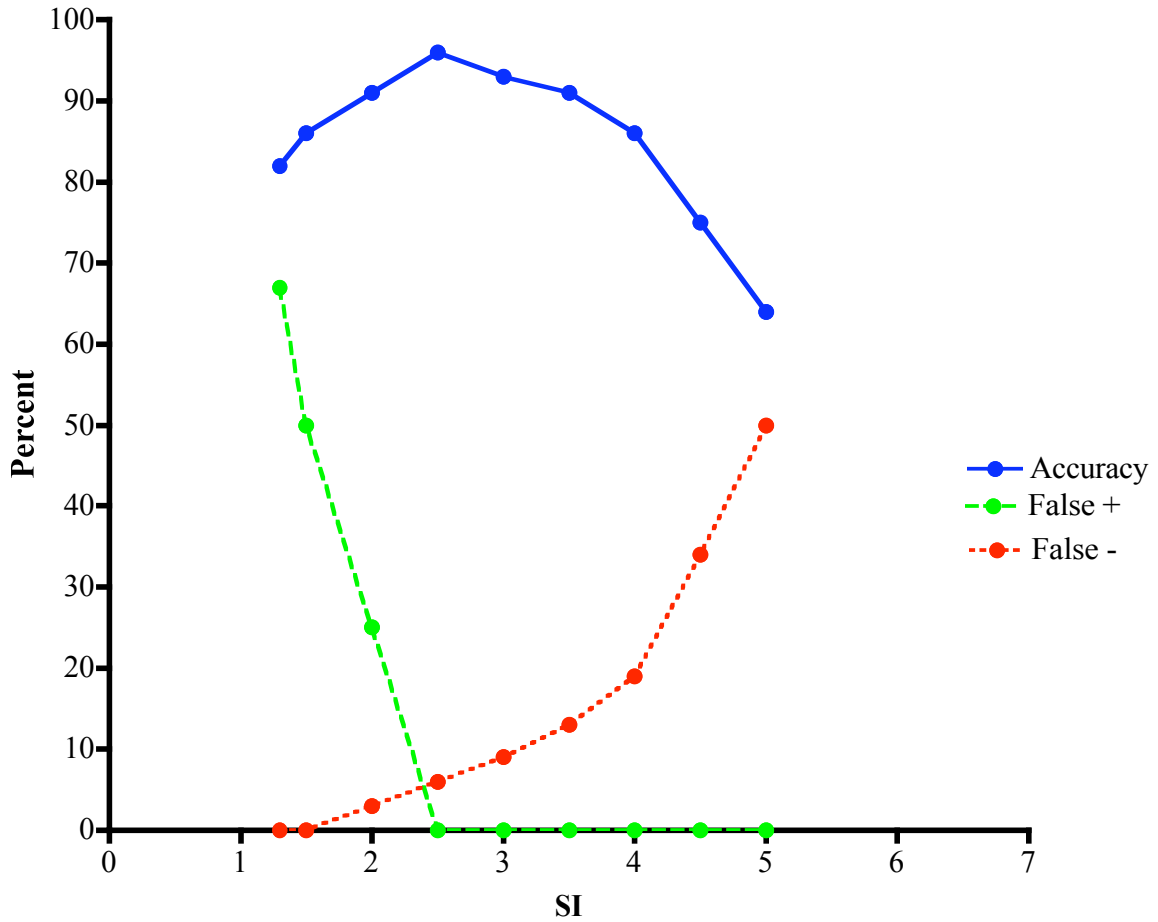
47 **1.1 Results of LLNA: DA Accuracy Analysis Using Alternative Decision Criteria**
48 **and Highest Maximum SI for the Outcome of Multiple Tests**

49 When combining multiple test results for a single substance by using the outcome of the test
50 with the highest maximum SI, the decision criterion of $SI \geq 3.0$ (used by the LLNA: DA
51 validation study team) to identify sensitizers yielded an accuracy of 93% (41/44), a
52 sensitivity of 91% (29/32), a specificity of 100% (12/12), a false positive rate of 0% (0/12),
53 and a false negative rate of 9% (3/32) (**Table E-1**). The decision criteria using higher SI
54 values, $SI \geq 3.5$ to $SI \geq 5.0$, decreased performance except for the specificity and the false
55 positive rate, which remained at 100% (12/12) and 0% (0/12), respectively (**Figure E-1** and
56 **Table E-1**). The lower SI criterion, $SI \geq 2.5$, increased accuracy to 96% (42/44) and
57 sensitivity to 94% (30/32), while the specificity and the false positive rate remained the same
58 at 100% (12/12) and 0% (0/12), respectively. Further, the false negative rate decreased to 6%

59 (2/32) at $SI \geq 2.5$. At an even lower SI criterion, $SI \geq 1.3$, accuracy was 82% (36/44) and
60 sensitivity was 100% (32/32), while the specificity was low (33% [4/12]) and the false
61 positive rate was high (67% [8/12]). Further, the false negative rate decreased to 0% (0/32) at
62 $SI \geq 1.3$. The use of analysis of variance (ANOVA) and summary statistics (i.e., mean ATP
63 measurement of treated groups $\geq 95\%$ confidence interval (CI) of the control group, or ≥ 2 or
64 ≥ 3 standard deviation [SD] from the control group mean), yielded accuracy values of 75% to
65 84%, with sensitivity values of 88% to 100%, and false negative rates of 0 to 13%. The
66 specificity for these criteria ranged from 8% to 58% and the false positive rates were 42% to
67 92%. Of these alternative decision criteria, the best overall performance for the approach
68 using the highest maximum SI for the substances with more than one test was achieved using
69 an $SI \geq 2.5$, as summarized above. Using an $SI \geq 2.5$, however, incorrectly classified 2-
70 mercaptobenzothiazole, a commonly know skin sensitizer.

71

71 **Figure E-1 Performance of the LLNA: DA with SI Compared to the Traditional**
 72 **LLNA Using the Highest Maximum SI for Substances with Multiple**
 73 **Tests**



74

75 As compared to traditional LLNA results, the lines show the change in performance characteristics
 76 for the LLNA: DA with the SI cutoff used to identify sensitizers. This analysis used LLNA: DA and
 77 traditional LLNA results for 44 substances (32 traditional LLNA sensitizers and 12 traditional LLNA
 78 nonsensitizers). For the 14 substances with multiple test results, the results for each substance were
 79 combined by using the outcome for the test with the highest maximum SI value. The solid line shows
 80 accuracy, the dashed line shows the false positive rate, and the dotted line shows the false negative
 81 rate.

82 **Table E-1 Performance of the LLNA: DA Compared with the Traditional LLNA Using Alternative Decision Criteria to**
 83 **Identify Sensitizers Based on the Highest 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	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
SI ≥ 3.0	44	93	41/44	91	29/32	100	12/12	0	0/12	9	3/32	100	29/29	80	12/15
SI ≥ 2.5	44	96	42/44	94	30/32	100	12/12	0	0/12	6	2/32	100	30/30	86	12/14
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 ≥ 1.5	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

84 Bolded text indicates the decision criterion chosen by the LLNA: DA validation study team; Italicized text indicates the single decision criterion that had an overall increased performance in predicting
 85 skin sensitization potential when compared to the traditional LLNA.

86 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;
 87 No. = number; SD = standard deviation; SI = stimulation index

88 ¹N = Number of substances included in this analysis.

89 ²The proportion on which the percentage calculation is based.

90 ³Analysis 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
 91 statistical analyses. For analysis of variance, significance at *p* < 0.05 was further tested by Dunnett's test.

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

93
94

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

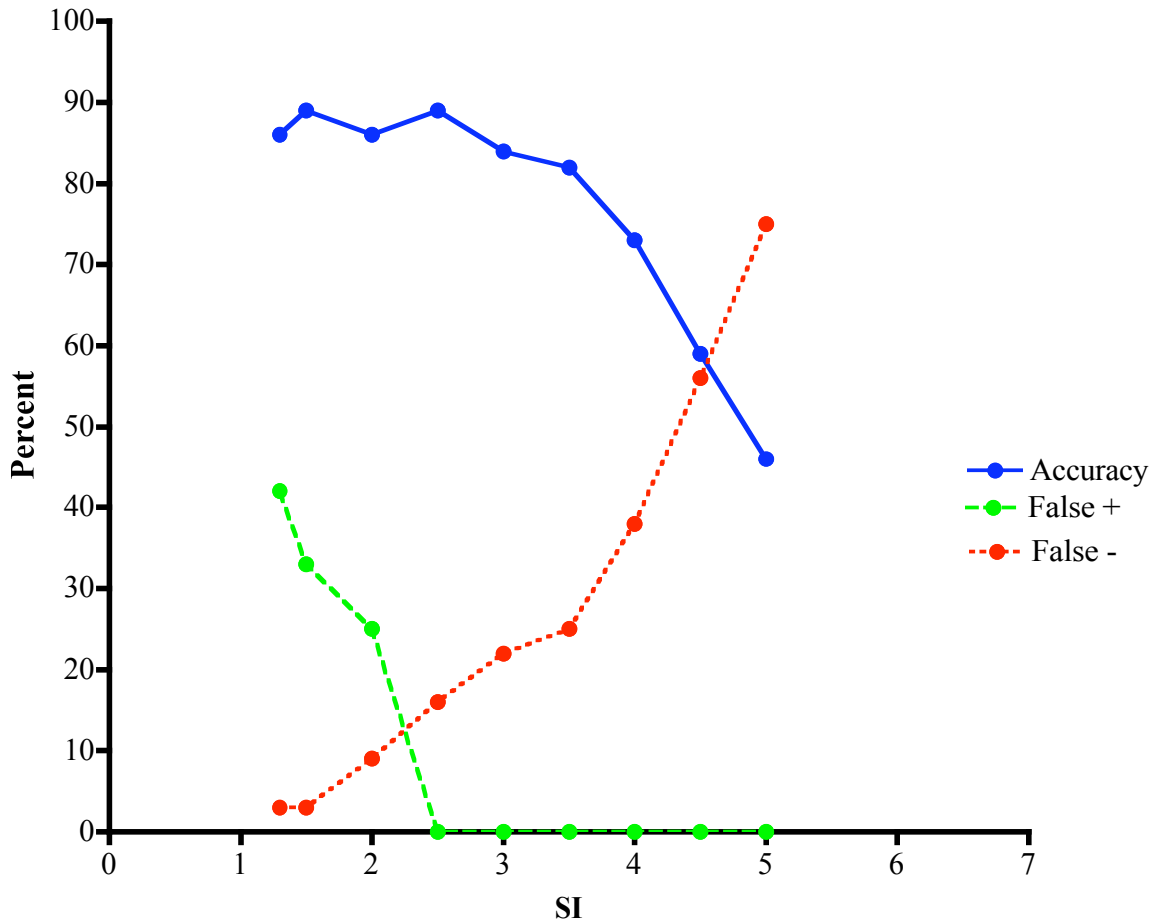
⁶ The mean ATP of at least one treatment group was greater than 3 SD from the mean ATP of the vehicle control group.

95 **1.2 Results of LLNA: DA Accuracy Analysis Using Alternative Decision Criteria**
96 **and Lowest Maximum SI for the Outcome of Multiple Tests**

97 When combining multiple test results for a single substance using the outcome of the test
98 with the lowest maximum SI to identify sensitizers, the decision criterion of $SI \geq 3.0$ (used by
99 the LLNA: DA validation study team) yielded an accuracy of 84% (37/44), a sensitivity of
100 78% (25/32), a specificity of 100% (12/12), a false positive rate of 0% (0/12), and a false
101 negative rate of 22% (7/32) (**Table E-2**). The decision criteria using higher SI values,
102 $SI \geq 3.5$ to $SI \geq 5.0$, decreased performance except for the specificity and the false positive
103 rate, which remained at 100% (12/12) and 0% (0/12), respectively (**Figure E-2** and **Table E-**
104 **2**). At $SI \geq 5.0$, accuracy decreased to 46% (20/44) and the false negative rate increased to
105 75% (24/32). Use of a lower SI at $SI \geq 2.5$ increased accuracy to 89% (39/44) and sensitivity
106 to 84% (27/32), while the specificity and false positive rate remained the same at 100%
107 (12/12) and 0% (0/12), respectively. Further, the false negative rate decreased to 16% (5/32)
108 at $SI \geq 2.5$. At an even lower SI criterion, $SI \geq 1.3$, accuracy was decreased to 86% (38/44)
109 but the sensitivity increased to 97% (31/32), while the specificity was 58% (7/12) and the
110 false positive rate was 42% (5/12). Further, the false negative rate decreased to 3% (1/32) at
111 $SI \geq 1.3$. Use of a statistical test (i.e., ANOVA or *t*-test) and summary statistics (i.e., mean
112 ATP measurements of treated groups $\geq 95\%$ CI of the control group, or ≥ 2 or ≥ 3 SD from the
113 control group mean), yielded accuracy values of 77 to 82%, with sensitivity values of 84 to
114 97%, and false negative rates of 3 to 16%. Both the specificity and false positive rate for
115 these criteria ranged from 42 to 58%. Of these alternative decision criteria, the best overall
116 performance for the approach using the lowest maximum SI for the substances with more
117 than one test was achieved using $SI \geq 2.5$, as summarized above.

118

118 **Figure E-2 Performance of the LLNA: DA with SI Compared to the Traditional**
 119 **LLNA Using the Lowest Maximum SI for Substances with Multiple Tests**



120

121 As compared to traditional LLNA results, the lines show the change in performance characteristics
 122 for the LLNA: DA with the SI cutoff used to identify sensitizers. This analysis used LLNA: DA and
 123 traditional LLNA results for 44 substances (32 traditional LLNA sensitizers and 12 traditional LLNA
 124 nonsensitizers). For the 14 substances with multiple test results, the results for each substance were
 125 combined by using the outcome for the test with the lowest maximum SI value. The solid line shows
 126 accuracy, the dashed line shows the false positive rate, and the dotted line shows the false negative
 127 rate.

128 **Table E-2 Performance of the LLNA: DA Compared with the Traditional LLNA Using Alternative Decision Criteria to**
 129 **Identify Sensitizers 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
SI ≥ 3.0	44	84	37/44	78	25/32	100	12/12	0	0/12	22	7/32	100	25/25	63	12/19
<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 ≥ 1.5	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

130 Bolded text indicates the decision criterion chosen by the LLNA: DA validation study team; Italicized text indicates the single decision criterion that had an overall increased performance in predicting
 131 skin sensitization potential when compared to the traditional LLNA.

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

134 ¹N = Number of substances included in this analysis.

135 ²The proportion on which the percentage calculation is based.

136 ³Analysis 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
 137 statistical analyses. For analysis of variance, significance at *p* < 0.05 was further tested by Dunnett's test.

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

139
140
141

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.

.

142 **2.0 Discordant Results for Accuracy Analysis of Alternative Decision** 143 **Criteria**

144 As mentioned above, for the 14 substances with multiple test results using the decision
145 criteria of $SI \geq 2.5$ to identify sensitizers (based on the test with the highest maximum SI)
146 yielded the best overall performance among the alternative decision criteria evaluated. When
147 compared to the traditional LLNA, 2-mercaptobenzothiazole, a well-known skin sensitizer
148 was misclassified as a nonsensitizer (**Table E-3**).

149 **2.1 Discordant Results Using Alternative Decision Criteria and Highest** 150 **Maximum SI Outcome for Multiple Tests**

151 Using the decision criterion of $SI \geq 3.0$ to identify sensitizers and the test with the highest
152 maximum SI as the representative result for substances with multiple tests yielded three
153 discordant substances (i.e., 3-aminophenol, 2-mercaptobenzothiazole, and methyl
154 methacrylate), all misclassified as nonsensitizers by the LLNA: DA.

155 **Table E-3** shows how the number and identity of discordant substances changes with the
156 alternate decision criteria when using the test with the highest maximum SI to represent the
157 outcome for substances with multiple tests. Using an SI cutoff lower than three to identify
158 sensitizers, such as $SI \geq 2.0$, yielded four discordant substances: chlorobenzene, hexane, and
159 salicylic acid were misclassified as sensitizers and methyl methacrylate was misclassified as
160 a nonsensitizer. Using an even lower SI cutoff of $SI \geq 1.3$ to identify sensitizers, yielded five
161 additional discordant substances that were all misclassified as sensitizers (i.e., 1-
162 bromobutane, dimethyl isophthalate, isopropanol, methyl salicylate, and nickel [II] chloride).
163 Increasing the SI cutoff to values greater than three, increased the number of sensitizers that
164 were misclassified as nonsensitizers. At $SI \geq 4.0$, six traditional LLNA sensitizers were
165 misclassified as nonsensitizers by the LLNA: DA while at $SI \geq 5.0$, 17 sensitizers were
166 classified as nonsensitizers (**Table E-3**).

167 Use of a statistical test (i.e., ANOVA or *t*-test) or summary statistics (i.e., $\geq 95\%$ CI, ≥ 2 SD,
168 or ≥ 3 SD) tended to misclassify nonsensitizers in the traditional LLNA as sensitizers in the
169 LLNA: DA. Using ANOVA or *t*-test to identify sensitizers misclassified five nonsensitizers
170 (i.e., 1-bromobutane, chlorobenzene, hexane, salicylic acid, and sulfanilamide) as sensitizers

171 and two sensitizers (i.e., 2-mercaptobenzothiazole and methyl methacrylate) as
172 nonsensitizers. Using treatment group ATP measurement with ≥ 2 SD or ≥ 3 SD of the vehicle
173 control mean or a $\geq 95\%$ CI of the vehicle control mean, all misclassified the following six
174 traditional LLNA nonsensitizers as sensitizers: 1-bromobutane, chlorobenzene, hexane,
175 isopropanol, nickel (II) chloride, and propylparaben. The $\geq 95\%$ CI of the vehicle control
176 mean misclassified four additional nonsensitizers (i.e., diethyl phthalate, dimethyl
177 isophthalate, lactic acid, and methyl salicylate) as sensitizers. In addition, ≥ 2 SD or ≥ 3 SD of
178 the vehicle control mean commonly misclassified three sensitizers as nonsensitizers (i.e.,
179 ethyl acrylate, methyl methacrylate, and propyl gallate).

180 Thirteen of the 22 ICCVAM-recommended LLNA performance standards reference
181 substances (ICCVAM 2009) tested in the LLNA: DA were discordant for the analysis of
182 alternate decision criteria using the test with the highest maximum SI to represent substances
183 with multiple tests (**Table E-3**) when compared to the traditional LLNA. Six nonsensitizers
184 in the traditional LLNA (i.e., chlorobenzene, isopropanol, lactic acid, methyl salicylate,
185 nickel [II] chloride, and salicylic acid) were misclassified by some criteria in the LLNA: DA
186 as a sensitizers, and seven sensitizers in the traditional LLNA (i.e., citral, ethylene glycol
187 dimethacrylate, imidazolidinyl urea, 2-mercaptobenzothiazole, methyl methacrylate, phenyl
188 benzoate, and sodium lauryl sulfate) were misclassified as nonsensitizers by some criteria
189 when tested in the LLNA: DA.

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

Discordant Substance ¹	Alternate 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 ≥ 1.5	SI ≥ 1.3
3-Aminophenol (3.2%)					-	-	-	-	-				
p-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	-				-	-	-	-	-	-			

Discordant Substance ¹	Alternate 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 ≥ 1.5	SI ≥ 1.3
(1.7%)													
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%)					-								

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

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

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

196 ³Analysis of variance assessed difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at
 197 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
 198 Dunnett’s test.

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

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

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

202 **2.2 Discordant Results Using Alternative Decision Criteria and Lowest Maximum**
203 **SI Outcome for Multiple Tests**

204 As mentioned above, for the substances with multiple tests, using the decision criterion of
205 $SI \geq 2.5$ to identify sensitizers (based on the test with the lowest maximum SI) yielded the
206 best overall performance for the LLNA: DA when compared to the traditional LLNA. This
207 decision criterion yielded five discordant substances; all five were sensitizers in the
208 traditional LLNA but were misclassified as nonsensitizers in the LLNA: DA (i.e., 3-
209 aminophenol, cobalt chloride, 2-mercaptobenzothiazole, methyl methacrylate, and nickel [II]
210 sulfate hexahydrate) (**Table E-4**).

211 **Table E-4** shows how the number and identity of discordant substances changes with the
212 alternate decision criteria when using the test with the lowest maximum SI as the
213 representative result for substances with multiple tests. Using an SI cutoff less than three,
214 $SI \geq 2.0$, to identify sensitizers yielded six discordant substances. Three of the six discordant
215 substances (i.e., 3-aminophenol, methyl methacrylate, and nickel [II] sulfate hexahydrate)
216 were misclassified as nonsensitizers by the LLNA: DA compared to the traditional LLNA
217 and the remaining three (i.e., chlorobenzene, hexane, and salicylic acid) were misclassified as
218 sensitizers. Using an even lower SI to identify sensitizers, $SI \geq 1.3$, also yielded six discordant
219 substances. Chlorobenzene, hexane, and salicylic acid were still misclassified as sensitizers
220 and nickel (II) sulfate hexahydrate was still misclassified as a nonsensitizer by the LLNA:
221 DA compared to the traditional LLNA. In addition, 1-bromobutane and nickel (II) chloride
222 were also misclassified as sensitizers. Increasing the SI cutoff to values greater than three,
223 increased the number of sensitizers that were misclassified as nonsensitizers. At $SI \geq 4.0$, 12
224 sensitizers were misclassified as nonsensitizers while at $SI \geq 5.0$, 24 sensitizers were
225 misclassified as nonsensitizers (**Table E-4**). Using the test with the lowest maximum SI as
226 the result for substances with multiple tests caused even potent sensitizers to be misclassified
227 as nonsensitizers at the higher SI cutoffs. For instance, at $SI \geq 5.0$, 2,4-dinitrochlorobenzene
228 and glutaraldehyde were classified as nonsensitizers.

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

Discordant Substance ¹	Alternate 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 ≥ 1.5	SI ≥ 1.3
Abietic Acid (11.9%)					-	-	-						
3-Aminophenol (3.2%)					-	-	-	-	-	-	-		
p-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	Alternate 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 ≥ 1.5	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%)					-								

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

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

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

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

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

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

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

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

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

264