Annex VI

Accuracy Analyses Using Additional Approaches for Combining Multiple Test Results

This page intentionally left blank

1.0 LLNA: BrdU-ELISA Accuracy Analysis Using Alternative Decision Criteria and Alternate Methods for Combining Data for Substances Tested Multiple Times

This annex shows performance analyses for the LLNA: BrdU-ELISA using single alternative decision criteria and two different approaches for combining test results for the 18 substances with multiple LLNA: BrdU-ELISA tests:

1. The positive/negative outcome for each substance for each criterion was determined by the outcome of the test with the highest maximum 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.

Appendix C, Section 6.5 provides the results for the analysis when the most prevalent outcome for each criterion was used as the result for each substance that was tested multiple times.

1.1 Results of LLNA: BrdU-ELISA Accuracy Analysis Using Single Alternative Decision Criteria and the Highest 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 highest maximum SI, the decision criterion of SI ≥ 2.0 to identify sensitizers yielded an accuracy of 91% (39/43), a sensitivity of 91% (29/32), a specificity of 91% (10/11), a false positive rate of 9% (1/11), and a false negative rate of 9% (3/32) (**Table C-VI-1)**. SI ≥ 2.0 was the decision criterion used by the JSAAE interlaboratory validation study of the LLNA: BrdU-ELISA. The performance for the additional decision criteria is shown in **Table C-VI-1**. Over the range of SI cutoffs evaluated, the SI cutoffs of 1.5 and 1.9 to 2.5 produced the same accuracy as SI ≥ 2.0 (i.e., 91%). Accuracy at the optimum criterion of SI ≥ 1.6, identified in **Section 6.5** of the BRD, was the highest, at 93% (40/43). At SI ≥ 1.3, and at SI ≥ 3.0 and higher, the accuracy decreased to 88% (38/43) and from 86% (37/43) to 56% (24/43), respectively. SI ≥ 2.5 had the same sensitivity as SI ≥ 2.0 (91% [29/32]), but higher SI cutoffs decreased sensitivity (81% [26/32]) at SI ≥ 3.0 to 41% [13/32] at SI ≥ 5.0), increased specificity (100% [11/11] from SI ≥ 3.0 to SI ≥ 5.0), decreased the false positive rate (0% [0/11] at SI ≥ 3.0 and higher), and increased the false negative rate (19% [6/32] at SI ≥ 3.0 to 59% [19/32] at SI ≥ 5.0) (**Figure C-VI-1** and **Table C-VI-1)**. SI cutoffs lower than 2.0 increased sensitivity (94% [30/32] at SI ≥ 1.9 to 100% [32/32] at SI ≥ 1.6 to 1.3), decreased specificity (82% [9/11] at SI ≥ 1.9 to 55% [6/11] at SI ≥ 1.3), increased the false positive rate (18% [2/11] at SI ≥ 1.9 to 45% [5/11] at SI ≥ 1.3), and decreased the false negative rate (6% [2/32] at SI ≥ 1.9 to 0% [0/32] at SI ≥ 1.6 to 1.3). Use of ANOVA and summary statistics (i.e., mean absorbance values of treated groups ≥95% confidence interval [CI] of the control group mean, or ≥2 or ≥3 SD from the control group mean), yielded accuracy values of 86% (37/43) to 91% (39/43), with sensitivity values of 94% (30/32) to 100% (32/32), and false negative rates of 0% (0/32) to 6% (2/32). The specificity for these criteria ranged from 45% (5/11) to 82% (9/11) and the false positive rates were 18% (2/11) to 55% (6/11).

Table C-VI-1 Performance of the LLNA: BrdU-ELISA for 43 Substances Compared with the Traditional LLNA Using Alternative Decision Criteria to Identify Sensitizers and the Highest Maximum SI for Substances with Multiple Tests

Alternate Criterion	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate		Positive Predictivity		Negative Predictivity	
	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]
Statistics[2]	88	38/43	97	31/32	64	7/11	36	4/11	3	1/32	89	31/35	88	7/8
≥ 95% CI[3]	86	37/43	100	32/32	45	5/11	55	6/11	0	0/32	84	32/38	100	5/5
≥ 2 SD[4]	86	37/43	100	32/32	45	5/11	55	6/11	0	0/32	84	32/38	100	5/5
≥ 3 SD[5]	91	39/43	94	30/32	82	9/11	18	2/11	6	2/32	94	30/32	82	9/11
SI ≥ 5.0	56	24/43	41	13/32	100	11/11	0	0/11	59	19/32	100	13/13	37	11/30
SI ≥ 4.5	67	29/43	56	18/32	100	11/11	0	0/11	44	14/32	100	18/18	44	11/25
SI ≥ 4.0	72	31/43	63	20/32	100	11/11	0	0/11	38	12/32	100	20/20	48	11/23
SI ≥ 3.5	79	34/43	72	23/32	100	11/11	0	0/11	28	9/32	100	23/23	55	11/20
SI ≥ 3.0	86	37/43	81	26/32	100	11/11	0	0/11	19	6/32	100	26/26	65	11/17
SI ≥ 2.5	91	39/43	91	29/32	91	10/11	9	1/11	9	3/32	97	29/30	77	10/13
SI ≥ 2.0	91	39/43	91	29/32	91	10/11	9	1/11	9	3/32	97	29/30	77	10/13
SI ≥ 1.9	91	39/43	94	30/32	82	9/11	18	2/11	6	2/32	94	30/32	82	9/11
SI ≥ 1.6	93	40/43	100	32/32	73	8/11	27	3/11	0	0/32	91	32/35	100	8/8
SI ≥ 1.5	91	39/43	100	32/32	64	7/11	36	4/11	0	0/32	89	32/36	100	7/7
SI ≥ 1.3	88	38/43	100	32/32	55	6/11	45	5/11	0	0/32	86	32/37	100	6/6

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine; No. = number; SD = standard deviation; SI = stimulation index.

[1] The proportion on which the percentage calculation is based.

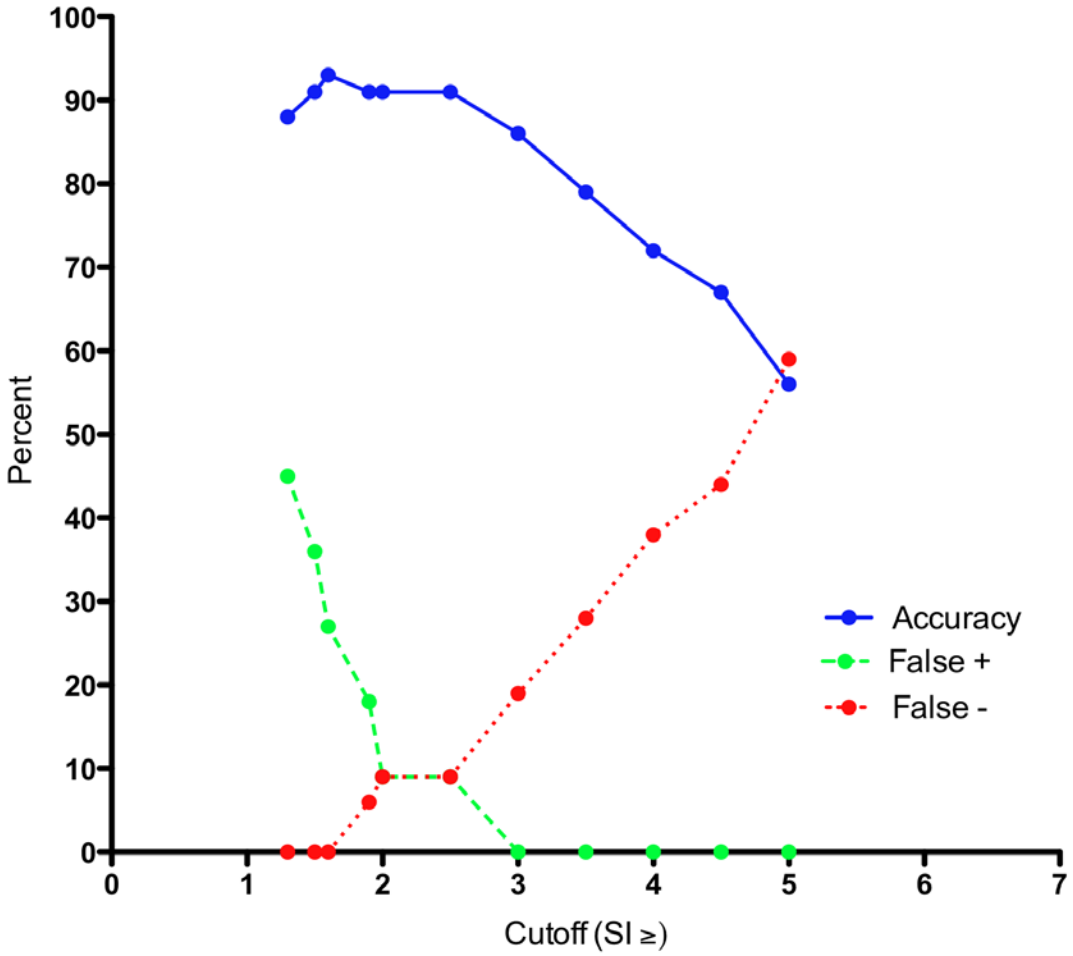
[2] Analysis of variance for difference of group means when substances were tested at multiple doses or *t*-test when substances were tested at one dose. The absorbance data were log-transformed prior to analysis of variance. Significance at $p < 0.05$ was further tested by Dunnett's test.

[3] The mean absorbance of at least one treatment group was outside the 95% confidence interval for the mean absorbance of the vehicle control group.

[4] The mean absorbance of at least one treatment group was greater than 3 SD from the mean absorbance of the vehicle control group.

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

Figure C-VI-1 Performance of the LLNA: BrdU-ELISA for 43 Substances with SI Compared to the Traditional LLNA Using the Highest Maximum SI for Substances with Multiple Tests



Abbreviations: LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine; SI = stimulation index.

The highest accuracy and lowest false negative rate for the approach using the highest maximum SI for the substances with more than one test, was achieved using an SI ≥ 1.6, the optimum criterion identified in **Section 6.5** of the BRD. The accuracy for SI ≥ 1.6 was 93% (40/43), with sensitivity of 100% (32/32), specificity of 73% (8/11), a false positive rate of 27% (3/11), and a false negative rate of 0% (0/32). However, using an SI ≥ 1.6 incorrectly classified lactic acid and isopropanol, two of the ICCVAM performance standards reference substances, as sensitizers. Use of mean absorbance values of treated groups ≥95% CI of the control group mean, or ≥2 SD from the control group mean, to identify sensitizers also produced the false negative rates as low as the SI cutoffs of 1.6 to 1.3 (0% [0/32]), with a slightly lower accuracy of 86% (37/43) to 88% (38/43), and a higher false positive rate of 55% (6/11). These criteria also incorrectly classified lactic acid and isopropanol, as well as methyl salicylate, another ICCVAM performance standards reference substance, as sensitizers. The

lowest false positive rates (0% [0/11]) were produced by SI cutoffs of 3.0 to 5.0; however, the false negative rates at those cutoffs were 19% (6/32) to 59% (19/32).

As compared to traditional LLNA results, the lines show the change in performance characteristics for the LLNA: BrdU-ELISA with the SI cutoff used to identify sensitizers. This analysis used LLNA: BrdU-ELISA and traditional LLNA results for 43 substances (32 sensitizers and 11 nonsensitizers based on traditional LLNA results). For the 18 substances with multiple test results, the results for each substance were combined by using the outcome for 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.

1.2 Results of LLNA: BrdU-ELISA 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, the decision criterion of SI ≥ 2.0 to identify sensitizers for these 43 substances yielded an accuracy of 86% (37/43), a sensitivity of 81% (26/32), a specificity of 100% (11/11), a false positive rate of 0% (0/11), and a false negative rate of 19% (6/32) (**Table C-VI-2)**. SI ≥ 2.0 was the decision criterion used by the JSAAE interlaboratory validation study of the LLNA: BrdU-ELISA. The performance for the additional decision criteria is shown in **Table C-VI-2**.

Over the range of SI cutoffs evaluated, increasing the SI cutoff compared with SI ≥ 2.0 decreased accuracy (79% [34/43] at SI ≥ 2.5 to 40% [21/43] at SI ≥ 5.0), decreased sensitivity (72% [23/32] at SI ≥ 2.5 to 19% [6/32] at SI ≥ 5.0), produced the same specificity (100% [11/11] up to SI ≥ 5.0), produced the same false positive rate (0% [0/11] up to SI ≥ 5.0), and increased the false negative rate (28% [9/32] at SI ≥ 2.5 to 81% [26/32] at SI ≥ 5.0) (**Figure C-VI-2** and **Table C-VI-2**). SI cutoffs lower than 2.0 increased sensitivity (88% [28/32] at SI ≥ 1.9 to 100% [32/32] at SI ≥ 1.3), decreased specificity (100% [11/11] at SI ≥ 1.9 to 73% [8/11] at SI ≥ 1.3), increased the false positive rate (0% [0/11] at SI ≥ 1.9 to 27% [3/11] at SI ≥ 1.3), and decreased the false negative rate (13% [4/32] at SI ≥ 1.9 to 0% [0/32] at SI ≥ 1.3). Use of ANOVA and summary statistics (i.e., mean absorbance values of treated groups ≥95% CI of the control group mean, or ≥2 or 3 SD from the control group mean), yielded accuracy of 88% (38/43) to 93% (40/43), with sensitivity values of 88% (28/32) to 97% (31/32), and false negative rates of 3% (1/32) to 13% (4/32). The specificity for these criteria ranged from 64% (7/11) to 100% (11/11) and the false positive rates were 0% (0/11) to 36% (4/11).

The highest accuracy and lowest false negative rate for the approach using the lowest maximum SI for the substances with more than one test was achieved using an SI ≥ 1.3 and mean absorbance values of treated groups ≥2 SD of the control group mean. Both criteria yielded an accuracy of 93% (40/43). The false negative rates were 0% (0/32) at SI ≥ 1.3 and 3% (1/32) at ≥2 SD. The sensitivity was 100% (32/32) at SI ≥ 1.3 and 97% (31/32) at ≥2 SD. The specificity was 73% (8/11) at SI ≥ 1.3 and 82% (9/11) at ≥2 SD. The false positive rate was 27% (3/11) at SI ≥ 1.3 and 18% (2/11) at ≥2 SD. However, SI ≥ 1.3 incorrectly classified lactic acid and methyl salicylate, two of the ICCVAM performance standards reference substances, as sensitizers. Mean absorbance values of treated groups ≥2 SD from the control group mean incorrectly classified lactic acid as a sensitizer. The lowest false positive rate (0% [0/11]) was produced by SI cutoffs of 1.9 to 5.0 and mean absorbance values of treated groups ≥3 SD; however, the false negative rates at those cutoffs were 13% (4/32) to 81% (26/32). Of those cutoffs, SI ≥ 1.9 and mean absorbance values of treated groups ≥3 SD produced the highest accuracy, 91% (39/43), and the lowest false negative rates, 13% (4/32).

Table C-VI-2 Performance of the LLNA: BrdU-ELISA for 43 Substances Compared with the Traditional LLNA Using Alternative Decision Criteria to Identify Sensitizers and the Lowest Maximum SI for Substances with Multiple Tests

Alternate Criterion	Accuracy		Sensitivity		Specificity		False Positive Rate		False Negative Rate		Positive Predictivity		Negative Predictivity	
	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]	%	No.[1]
Statistics[2]	88	38/43	91	29/32	82	9/11	18	2/11	9	3/32	94	29/31	75	9/12
≥ 95% CI[3]	88	38/43	97	31/32	64	7/11	36	4/11	3	1/32	89	31/35	88	7/8
≥ 2 SD[4]	93	40/43	97	31/32	82	9/11	18	2/11	3	1/32	94	31/33	90	9/10
≥ 3 SD[5]	91	39/43	88	28/32	100	11/11	0	0/11	13	4/32	100	28/28	73	11/15
SI ≥ 5.0	40	17/43	19	6/32	100	11/11	0	0/11	81	26/32	100	6/6	30	11/37
SI ≥ 4.5	44	19/43	25	8/32	100	11/11	0	0/11	75	24/32	100	8/8	31	11/35
SI ≥ 4.0	49	21/43	31	10/32	100	11/11	0	0/11	69	22/32	100	10/10	33	11/33
SI ≥ 3.5	56	24/43	41	13/32	100	11/11	0	0/11	59	19/32	100	13/13	37	11/30
SI ≥ 3.0	67	29/43	56	18/32	100	11/11	0	0/11	44	14/32	100	18/18	44	11/25
SI ≥ 2.5	79	34/43	72	23/32	100	11/11	0	0/11	28	9/32	100	23/23	55	11/20
SI ≥ 2.0	86	37/43	81	26/32	100	11/11	0	0/11	19	6/32	100	26/26	65	11/17
SI ≥ 1.9	91	39/43	88	28/32	100	11/11	0	0/11	13	4/32	100	28/28	73	11/15
SI ≥ 1.6	91	39/43	94	30/32	82	9/11	18	2/11	6	2/32	94	30/32	82	9/11
SI ≥ 1.5	91	39/43	94	30/32	82	9/11	18	2/11	6	2/32	94	30/32	82	9/11
SI ≥ 1.3	93	40/43	100	32/32	73	8/11	27	3/11	0	0/32	91	32/35	100	8/8

Abbreviations: CI = confidence interval; LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine; No. = number; SD = standard deviation; SI = stimulation index.

[1] The proportion on which the percentage calculation is based.

[2] 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 absorbance data were log-transformed prior to analysis of variance. Significance at $p < 0.05$ was further tested by Dunnett's test.

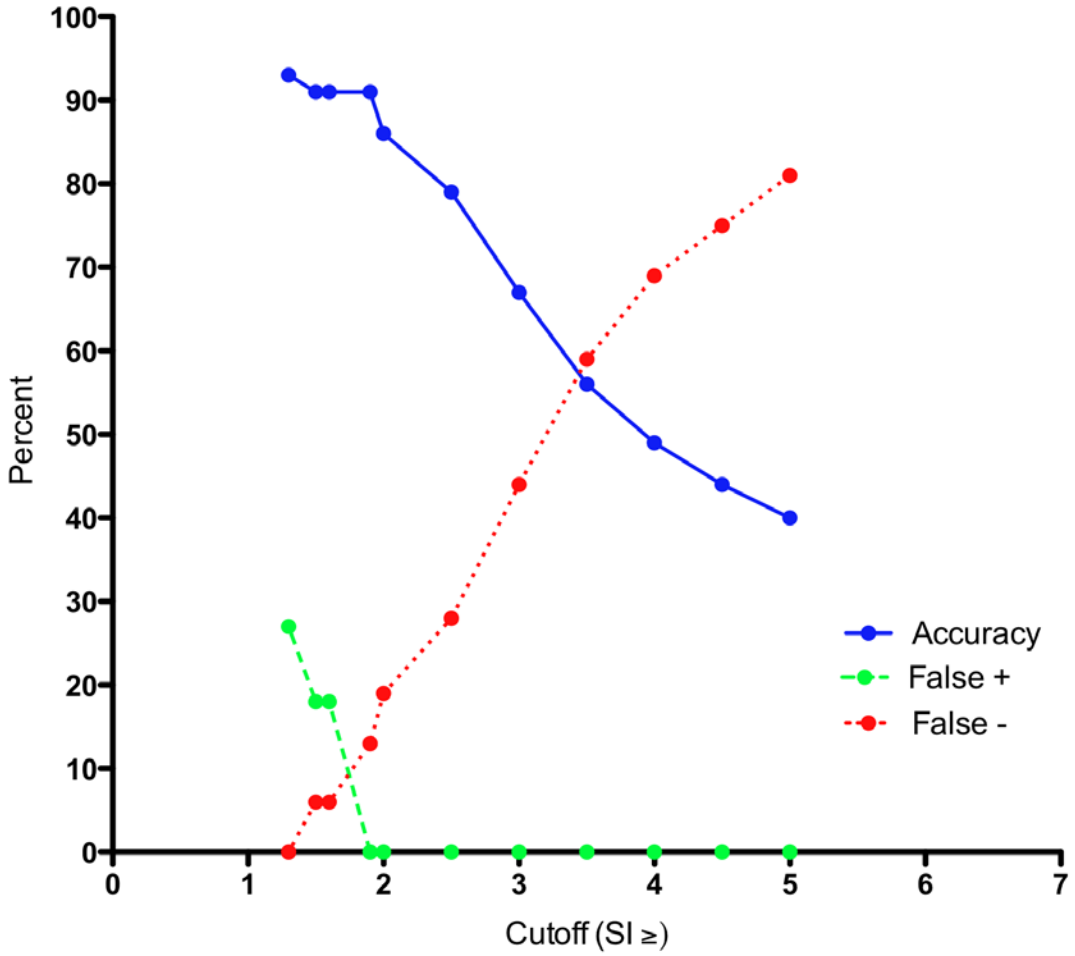
[3] The mean absorbance of at least one treatment group was outside the 95% confidence interval for the mean absorbance of the vehicle control group.

[4] The mean absorbance of at least one treatment group was greater than 3 SD from the mean absorbance of the vehicle control group.

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

.

Figure C-VI-2 Performance of the LLNA: BrdU-ELISA for 43 Substances with SI Compared to the Traditional LLNA Using the Lowest Maximum SI for Substances with Multiple Tests



Abbreviations: LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine; SI = stimulation index.

As compared to traditional LLNA results, the lines show the change in performance characteristics for the LLNA: BrdU-ELISA with the SI cutoff used to identify sensitizers. This analysis used LLNA: BrdU-ELISA and traditional LLNA results for 43 substances (32 sensitizers and 11 nonsensitizers based on traditional LLNA results). For the 18 substances with multiple test results, the results for each substance were combined by using the outcome for 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.

2.0 Discordant Results for Accuracy Analysis of Alternative Decision Criteria

Using the decision criteria of SI ≥ 2.0 to identify sensitizers and the most prevalent outcome for the substances with multiple tests, the LLNA: BrdU-ELISA outcomes yielded two discordant substances (2-mercaptobenzothiazole and imidazolidinyl urea) compared with the traditional LLNA (**Table C-5**). As indicated in **Appendix C, Section 6.4.1**, these substances were misclassified as nonsensitizers when compared to the traditional LLNA, which classified them as sensitizers.

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

Using the decision criteria of SI ≥ 2.0 to identify sensitizers and the test with the highest maximum SI as the result for substances with multiple tests, yielded two additional discordant substances: isopropanol and lactic acid, which were misclassified as sensitizers.

Table C-VI-3 shows how the number and identity of discordant substances change with the alternative decision criteria when using the test with the highest maximum SI as the result for substances with multiple tests. Using an SI cutoff less than 2.0 increased the number of traditional LLNA nonsensitizers that were misclassified as sensitizers. SI ≥ 1.3 yielded the highest number (five) of discordant substances that were misclassified as sensitizers (hexane, isopropanol, lactic acid, methyl salicylate, and propylene glycol). Increasing the SI cutoff to values greater than 2.0 increased the number of sensitizers that were misclassified as nonsensitizers. At SI ≥2.5, three sensitizers were misclassified as nonsensitizers while, at SI ≥ 5.0, 19 sensitizers were classified as nonsensitizers (**Table C-VI-3)**. At SI ≥ 2.0, two nonsensitizers were misclassified as sensitizers. Increasing the SI cutoff to values greater than 2.0 decreased the number of nonsensitizers classified as sensitizers. At SI ≥ 2.5, one nonsensitizer was misclassified as a sensitizer. At SI ≥ 3.0 and higher, no nonsensitizers were classified as sensitizers.

Use of a statistical test (i.e., ANOVA or *t*-test) or summary statistics (i.e., ≥95% CI, ≥2 SD, or ≥3 SD) tended to misclassify more nonsensitizers than sensitizers. Using ANOVA or a *t*-test to identify sensitizers misclassified one sensitizer (2-mercaptobenzothiazole) as a nonsensitizer and four nonsensitizers (glycerol, hexane, isopropanol, and lactic acid) as sensitizers. Using treatment group absorbance ≥95% CI or ≥2 SD of control group mean misclassified six nonsensitizers as sensitizers (glycerol, hexane, isopropanol, lactic acid, methyl salicylate, and propylene glycol). Using treatment group absorbance ≥3 SD of the control group mean misclassified two nonsensitizers as sensitizers (hexane and lactic acid) and two weak sensitizers as nonsensitizers (cinnamic alcohol and imidazolidinyl urea).

Eleven ICCVAM performance standards reference substances were discordant for the analysis of alternative decision criteria using the test with the highest maximum SI as the result for substances with multiple tests (**Table C-VI-3**). Eight traditional LLNA sensitizers (2-mercaptobenzothiazole, 5-chloro-2-methyl-4-isothiazolin-3-one, cinnamic alcohol, cobalt chloride, ethylene glycol dimethacrylate, imidazolidinyl urea, phenyl benzoate, and sodium lauryl sulfate) were misclassified by some criteria as nonsensitizers. Sodium lauryl sulfate, however, produces a false positive result in the traditional LLNA; it does not produce a sensitization reaction in humans or guinea pigs. Three nonsensitizers (isopropanol, lactic acid, and methyl salicylate) were misclassified as sensitizers by some criteria.

Table C-VI-3 Discordant Results for LLNA: BrdU-ELISA Using Alternative Decision Criteria Compared to the Traditional LLNA and the Highest Maximum SI for Substances with Multiple Tests

Discordant Substances[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≥1.9	SI≥1.6	SI≥1.5	SI≥1.3
Glycerol (-)	+	+	+												
Hexane (-)	+	+	+	+									+	+	+
Isopropanol[7] (-)	+	+	+								+	+	+	+	+
Lactic acid[7] (-)	+	+	+	+						+	+	+	+	+	+
Methyl salicylate[7] (-)		+	+												+
Propylene glycol (-)		+	+											+	+
2-Mercaptobenzothiazole[7] (1.7%)	-				-	-	-	-	-	-	-	-			
3-Aminophenol (3.2%)					-	-	-	-							
4-Chloroaniline (6.5%)					-	-	-	-	-						
MAPS (0.8%)					-	-	-								
CMI solution[7] (0.009%)					-										
Aniline (48%)					-	-	-	-	-	-					
Cinnamic alcohol[7] (21%)				-	-	-	-	-	-						
Cinnamic aldehyde (1.9%)					-	-	-								
Cobalt chloride[7] (0.6%)					-	-	-								
Ethyl acrylate (33%)					-										
EGDA[7] (28%)					-	-	-	-							
Hydroxycitronellal (24%)					-										
Imidazolidinyl urea[7] (24%)				-	-	-	-	-	-	-	-	-			
Isopropyl myristate (44%)					-	-									
Linalool (30%)					-										
Nickel sulfate (4.8%)					-										

Discordant Substances[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≥1.9	SI≥1.6	SI≥1.5	SI≥1.3
Phenyl benzoate[7] (14%)					-	-	-	-							
Sodium lauryl sulfate[7] (8.1%)					-	-	-	-	-						
trans-Cinnamaldehyde (1.4 %)					-	-									

Abbreviations: CI = confidence interval; CMI = 5-Chloro-2-methyl-4-isothiazolin-3-one; EGDA = ethylene glycol dimethacrylate; LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine (BrdU); MAPS = 4-methylaminophenol sulfate; SD = standard deviation; SI = stimulation index.

[1] Compared to the traditional LLNA. Traditional LLNA result in parentheses: "-" for nonsensitizers and EC3 (%) for sensitizers.

[2] LLNA: BrdU result shown: "+" if the decision criterion was met and "-" if the decision criterion was not met.

[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 absorbance data were log-transformed prior to analysis of variance. Significance at $p < 0.05$ was further tested by Dunnett's test.

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

[5] The mean absorbance of at least one treatment group was greater than 3 SD from the mean absorbance of the vehicle control group.

[6] The mean absorbance of at least one treatment group was greater than 2 SD from the mean absorbance of the vehicle control group.

[7] Reference substance from *Recommended Performance Standards: Murine Local Lymph Node Assay* (ICCVAM 2009; available: http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm).

The criteria that yielded the correct results for 2-mercaptobenzothiazole included the statistics ≥95% CI, ≥2 SD, or ≥3 SD; and SI ≥ 1.6 to SI ≥ 1.3. The criteria that yielded the correct results for 5-chloro-2-methyl-4-isothiazolin-3-one included all of the summary statistics and SI ≥ 4.5 to SI ≥ 1.3. The criteria that yielded the correct results for cinnamic alcohol included the summary statistics ANOVA, ≥95% CI, and ≥2 SD; and SI ≥ 2.5 to SI ≥ 1.3. The criteria that yielded the correct results for cobalt chloride included all of the summary statistics and SI ≥ 3.5 to SI ≥ 1.3. The criteria that yielded the correct results for ethylene glycol dimethacrylate and phenyl benzoate included all of the summary statistics and SI ≥ 3.0 to SI ≥ 1.3. The criteria that yielded the correct results for imidazolidinyl urea included the summary statistics ANOVA, ≥95% CI, and ≥2 SD; and SI ≥ 1.6 to SI ≥ 1.3. The criteria that yielded the correct results for isopropanol included treatment group absorbance ≥3 SD of vehicle control group mean and SI ≥ 2.5 and higher. The criteria that yielded the correct results for lactic acid included SI ≥ 3.0 and higher. All criteria yielded the correct results for methyl salicylate except for treatment group absorbance ≥95% CI or ≥2 SD of the control group mean, and SI ≥ 1.3.

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

Using the decision criteria of SI ≥ 2.0 to identify sensitizers and the most prevalent outcome for the substances with multiple tests, the LLNA: BrdU-ELISA outcomes yielded two discordant substances (2-mercaptobenzothiazole and imidazolidinyl urea) compared with the traditional LLNA (**Table C-5**). These substances were classified as nonsensitizers by the LLNA: BrdU-ELISA, while the traditional LLNA classified them as sensitizers. Using the test with the lowest maximum SI as the result for substances with multiple tests yielded six discordant substances at SI ≥ 2.0, including four additional sensitizers (cyclamen aldehyde, formaldehyde, hydroxycitronellal, and linalool), which were misclassified as nonsensitizers (**Table C-VI-4**). Linalool, however, is false positive in the traditional LLNA; it does not produce a sensitization reaction in humans (guinea pig data were not available).

Table C-VI-4 shows how the number and identity of discordant substances changed with the alternative decision criteria when using the test with the lowest maximum SI as the result for substances with multiple tests. Using an SI cutoff less than 2.0, SI ≥ 1.9 misclassified fewer (four) traditional LLNA sensitizers as nonsensitizers. SI ≥ 1.6 and lower increased the number of traditional LLNA nonsensitizers that were misclassified as sensitizers. SI ≥ 1.6 yielded two discordant substances that were misclassified as sensitizers (hexane and lactic acid), while SI ≥ 1.3 misclassified three nonsensitizers as sensitizers (hexane, lactic acid, and methyl salicylate). Increasing the SI cutoff to values greater than 2.0 increased the number of sensitizers that were misclassified as nonsensitizers. At SI ≥ 2.5, nine sensitizers were misclassified as nonsensitizers while, at SI ≥ 5.0, 26 sensitizers were classified as nonsensitizers (**Table C-VI-4)**. Although no nonsensitizers were misclassified as sensitizers at SI ≥ 1.9 and higher, lower SI cutoffs misclassified some nonsensitizers as sensitizers (two at SI ≥ 1.6 and SI ≥ 1.5 and three at SI ≥ 1.3).

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. At SI ≥ 4.5 and SI ≥ 5.0, 2,4-dinitrochlorobenzene, 4-methyl aminophenol sulfate, cobalt chloride, glutaraldehyde, and formaldehyde were classified as nonsensitizers. Glutaraldehyde was classified as a nonsensitizer at SI cutoffs as low as 2.5, and formaldehyde was classified as a nonsensitizer at SI cutoffs as low as 2.0.

Table C-VI-4 Discordant Results for LLNA: BrdU-ELISA Using Alternative Decision Criteria Compared to the Traditional LLNA and the Lowest Maximum SI for Substances with Multiple Tests

Discordant Substances[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 ≥1.9	SI ≥1.6	SI ≥1.5	SI ≥1.3
Glycerol (-)	+	+	+												
Hexane (-)	+	+											+	+	+
Lactic acid[7] (-)		+	+										+	+	+
Methyl salicylate[7] (-)		+													+
2-Mercaptobenzothiazole[7] (1.7%)	-				-	-	-	-	-	-	-	-			
2,4-Dinitrochlorobenzene (0.05%)[7]					-	-									
3-Aminophenol (3.2%)					-	-	-	-							
4-Chloroaniline (6.5%)					-	-	-	-	-						
MAPS (0.8%)					-	-	-								
CMI solution[7] (0.009%)					-										
Aniline (48%)					-	-	-	-	-	-					
Cinnamic alcohol[7] (21%)				-	-	-	-	-	-						
Cinnamic aldehyde (1.9%)					-	-	-								
Cobalt chloride[7] (0.6%)					-	-	-								
Cyclamen aldehyde (22%)					-	-	-	-	-	-	-				
Ethyl acrylate (33%)					-										
EGDA[7] (28%)					-	-	-	-							
Eugenol[7] (10%)					-	-	-	-							
Formaldehyde (0.5%)					-	-	-	-	-	-	-				
Glutaraldehyde (0.08%)					-	-	-	-	-	-					
Hexyl cinnamic aldehyde[7] (9.7%)					-	-	-	-	-						
Hydroxycitronellal (24%)	-			-	-	-	-	-	-	-	-	-	-	-	
Imidazolidinyl urea[7] (24%)				-	-	-	-	-	-	-	-	-			
Isoeugenol[7] (1.5%)					-	-	-	-	-	-					
Isopropyl myristate (44%)					-	-									
Linalool (30%)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	
Nickel sulfate (4.8%)					-	-	-	-	-						

Discordant Substances[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 ≥1.9**	**SI ≥1.6**	**SI ≥1.5**	**SI ≥1.3**
Phenyl benzoate[7] (14%)					-	-	-	-							
Sodium lauryl sulfate[7] (8.1%)					-	-	-	-	-						
trans-Cinnamaldehyde (1.4%)					-	-	-	-							

Abbreviations: CI = confidence interval; CMI = 5-Chloro-2-methyl-4-isothiazolin-3-one; EGDA = ethylene glycol dimethacrylate; LLNA = murine local lymph node assay; LLNA: BrdU-ELISA = murine local lymph node assay with enzyme-linked immunosorbent assay detection of bromodeoxyuridine (BrdU); MAPS = 4-methylaminophenol sulfate; SD = standard deviation; SI = stimulation index.

[1] Compared to the traditional LLNA. Traditional LLNA result in parentheses: "-" for nonsensitizers and EC3 (%) for sensitizers.

[2] LLNA: BrdU result shown: "+" if the decision criterion was met and "-" if the decision criterion was not met.

[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 absorbance data were log-transformed prior to analysis of variance. Significance at $p < 0.05$ was further tested by Dunnett's test.

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

[5] The mean absorbance of at least one treatment group was greater than 3 SD from the mean absorbance of the vehicle control group.

[6] The mean absorbance of at least one treatment group was greater than 2 SD from the mean absorbance of the vehicle control group.

[7] Reference substance from *Recommended Performance Standards: Murine Local Lymph Node Assay* (ICCVAM 2009; available: http://iccvam.niehs.nih.gov/methods/immunotox/llna_PerfStds.htm).

Use of a statistical test (i.e., ANOVA or *t*-test) or summary statistics (i.e., ≥95% CI, ≥2 SD, or 3 ≥SD) more often misclassified nonsensitizers than sensitizers (**Table C-VI-4**). Using ANOVA or *t*-tests to identify sensitizers misclassified three sensitizers (2-mercaptobenzothiazole, hydroxycitronellal, and linalool) as nonsensitizers and two nonsensitizers (glycerol and hexane) as sensitizers. Using treatment group absorbance ≥95% CI of the control group mean misclassified glycerol, hexane, lactic acid, and methyl salicylate as sensitizers. Using treatment group absorbance ≥2 SD of the control group mean misclassified glycerol and lactic acid as sensitizers. Using treatment group absorbance ≥3 SD of the control group mean misclassified three weak sensitizers as nonsensitizers (cinnamic alcohol, hydroxycitronellal, and imidazolidinyl urea). Linalool was classified as a nonsensitizer by all of the summary statistics, which is discordant with traditional LLNA results; however, linalool is false positive in the traditional LLNA. It does not produce a sensitization reaction in humans (guinea pig data were not available).

Fourteen ICCVAM performance standards reference substances were discordant for the analysis of alternative decision criteria using the test with the lowest maximum SI as the result for substances with multiple tests (**Table C-VI-4**). Three strong sensitizers, 2,4-dinitrochlorobenzene, 5-chloro-2-methyl-4-isothiazolin-3-one, and cobalt chloride were misclassified by some criteria as nonsensitizers. Nine additional sensitizers, 2-mercaptobenzothiazole, cinnamic alcohol, ethylene glycol dimethacrylate, eugenol, hexyl cinnamic aldehyde, imidazolidinyl urea, isoeugenol, phenyl benzoate, and sodium lauryl sulfated, were also misclassified as nonsensitizers by some criteria. Sodium lauryl sulfate, however, produces a false positive result in the traditional LLNA; it does not produce a sensitization reaction in humans or guinea pigs.

The criteria that yielded the correct results for 2-mercaptobenzothiazole included the summary statistics ≥95% CI, ≥2 SD, or ≥3 SD; and SI ≥ 1.6 to SI ≥ 1.3. The criteria that yielded the correct results for 2,4-dinitrochlorobenzene were all but the SI ≥ 4.5 to 5.0 criteria. The criteria that yielded the correct results for 5-chloro-2-methyl-4-isothiazolin-3-one were all but the SI ≥ 5.0 criterion. The criteria that yielded the correct results for cinnamic alcohol were all except ≥3 SD and SI ≥ 3.0 to SI ≥ 5.0. The criteria that yielded the correct results for cobalt chloride were all but the SI ≥ 5.0 to SI ≥ 4.0 criteria. The criteria that yielded the correct results for ethylene glycol dimethacrylate, eugenol, and phenyl benzoate were all but the SI ≥ 5.0 to SI ≥ 3.5 criteria. The criteria that yielded the correct results for hexyl cinnamic aldehyde were all except SI ≥ 5.0 to SI ≥ 3.0. The criteria that yielded the correct results for imidazolidinyl urea were the summary statistics ANOVA, ≥95% CI, or ≥2 SD; and SI ≥ 1.6 to SI ≥ 1.3 and SI ≥ 3.0 to SI ≥ 5.0. The criteria that yielded the correct results for isoeugenol were all except SI ≥ 5.0 to SI ≥ 2.5.

Two nonsensitizers, lactic acid and methyl salicylate, from the list of ICCVAM performance standards reference substances, were misclassified as sensitizers by some criteria. The criteria that yielded the correct results for lactic acid were all except for treatment group absorbance ≥95% CIand ≥2 SD of the control group mean, and SI ≥ 1.6 to SI ≥ 1.3. The criteria that yielded the correct results for methyl salicylate were all except for treatment group absorbance ≥95% CIof the control group mean and SI ≥ 1.3.

This page intentionally left blank