

9.0 OTHER SCIENTIFIC REPORTS AND REVIEWS

9.1 Reports in the Peer Reviewed Literature

In addition to the reports discussed in previous sections (**Sections 6.0** and **7.0**), additional HET-CAM study reports were identified during the literature review. In many of these reports, inadequate information on the substances tested (e.g., identity not specific) and/or the results obtained from the *in vitro* or *in vivo* studies (e.g., qualitative but not quantitative HET-CAM data, group mean but not individual *in vivo* animal scores) precluded their use in an assessment of the performance characteristics of HET-CAM as described in **Sections 6.0** and **7.0**. This section provides a summary of reports where sufficient information was not available to include them in the performance assessment as well as the summary conclusions of the reports used for the analyses described in **Sections 6.0** and **7.0**. In addition, where applicable, an explanation as why some data could or could not be used as part of the performance evaluation is provided.

9.1.1 Bagley et al. (1992)

Investigators from five chemical and pharmaceutical companies conducted an evaluation of five alternative ocular toxicity test methods, which had been used by these companies in a tiered-testing approach to evaluate eye irritation potential. The study evaluated 12 chemicals and 20 formulations (components of the formulations were not provided). In this study, the *in vitro* scores were calculated as IS(A) values. Comparative *in vivo* rabbit eye test results were obtained from concurrent studies conducted in accordance with the method described Draize et al. (1944), and *in vivo* test data was presented as MAS.

The correlation analyses described in the study compared IS(A) values with MAS values. This correlation yielded a Pearson's coefficient of 0.77 and Spearman's coefficient of 0.85. No additional analyses on the performance of the test method were provided.

Individual rabbit *in vivo* data was obtained for a subset of substances evaluated in this study. These data were used to assess the performance of the HET-CAM test method for detecting ocular corrosives and severe irritants based on the GHS (UN 2003), EPA (1996), or EU (2001) classification systems. The results of these analyses for the subset of substances are provided in **Section 6.0**.

9.1.2 Balls et al. (1995)

Under the auspices of the British Home Office and Directorate General XI of the European Commission, a validation study on proposed alternatives to the *in vivo* rabbit ocular toxicity test method was conducted. The goal of the evaluation was to identify at least one non-whole animal test method that could be proposed to regulatory authorities as a replacement for the currently accepted *in vivo* ocular toxicity test method. For the HET-CAM test method, a total of 52 substances were evaluated in 60 tests in two to four laboratories. Four test substances were evaluated at two different concentrations and two substances were evaluated at three different concentrations. The ocular irritancy potential of the test substances were ranked in terms of MMAS (which ranged from 0 to 108). The test substances evaluated in the validation study were classified as acids (4), acyl halide (1),

alcohols (9), aldehyde (1), alkali (1), esters (6), heterocyclics (3), hydrocarbons (2), inorganic chemicals (4), ketones (3), organophosphate (1), pesticides (5), surfactants (6), and miscellaneous (6). In this study, the *in vitro* scores were calculated as Q-Scores and S-Scores. *In vivo* data for 46 of the test substances, which were generated in compliance with OECD TG 405, were obtained from historical sources. *In vivo* rabbit eye data for 14 of the test substances were obtained from concurrent studies conducted in compliance with OECD TG 405.

The authors concluded that the correlations between HET-CAM *in vitro* and *in vivo* scores were generally poor to moderate, regardless of the physicochemical properties of the substances tested. A summary of the range of Pearson's and Spearman's correlation coefficients obtained in this study for the full set of substances as well as various subgroups are provided in **Table 9-1**.

Since the *in vivo* test results were expressed as MMAS, the data provided in this report could not be used to evaluate the accuracy of HET-CAM for detecting ocular corrosives and severe irritants based on the GHS (UN 2003), EPA (1996), or EU (2001) classification systems. However, in response to a request from NICEATM, ECVAM forwarded mean HET-CAM scores from each testing laboratory. Raw *in vivo* data were obtained from ECETOC (ECETOC 1998). These data were used in the performance assessment of the HET-CAM test method described in **Section 6.0** and **Section 7.0**.

9.1.3 Blein et al. (1991)

A multicenter study of alternative ocular toxicity test methods was conducted under Oeuvre Pour l'Assistance aux Animaux de Laboratoire (OPAL). The study evaluated 40 substances representing different chemical categories and ocular irritancies. In this study, the *in vitro* scores were calculated as IS(A) values. Comparative *in vivo* rabbit eye test results were obtained from concurrent studies conducted in accordance with Draize et al. (1944). The *in vivo* scores were segregated into three different irritancy classifications (mild, moderate, and extreme); the rationale for the *in vivo* decision criteria was not provided.

The investigators reported that the HET-CAM test method overpredicted the irritancy potential of test substances when they were tested undiluted, while *in vitro* studies conducted with 10-fold dilutions provided a better correlation with the *in vivo* rabbit ocular test results. Using a 10-fold dilution, the irritancy potentials of two substances (acetone and formaldehyde) were underestimated when compared to the *in vivo* classification.

HET-CAM data in this report were presented in graphical form and no attempt was made to extrapolate the graphically presented data to mean HET-CAM scores. Thus, the test substances could not be classified according to the classification system described in **Section 5.0** and were not used in the accuracy analysis described in **Section 6.0**.

Table 9-1 *In Vitro/In Vivo* Range of Correlations Reported in Balls et al. (1995)

Index Score	Pearson's Correlation	Spearman's Correlation
<i>Full set of test substances (11-49 depending on endpoint)</i>		
HET-CAM Q-Score	0.310-0.517	0.441-0.596
HET-CAM S-Score	0.060-0.332	0.018-0.340
HET-CAM Q-Score, with cut-off at 2	0.416-0.527	0.462-0.588
HET-CAM S-Score, with cut-off at 2	0.089-0.320	0.069-0.329
<i>Chemicals soluble in water (5-25 depending on endpoint)</i>		
HET-CAM Q-Score	0.314-0.758	0.327-0.681
HET-CAM S-Score	0.137-0.309	0.082-0.357
HET-CAM Q-Score, with cut-off at 2	0.185-0.364	0.309-0.480
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated
<i>Chemicals insoluble in water (4-12 depending on endpoint)</i>		
HET-CAM Q-Score	0.232-0.445	0.345-0.688
HET-CAM S-Score	-0.922-0.716	-0.971-0.738
HET-CAM Q-Score, with cut-off at 2	0.370-0.609	0.396-0.651
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated
<i>Surfactants (12)</i>		
HET-CAM Q-Score	0.448-0.847	0.596-0.839
HET-CAM S-Score	Not Evaluated	Not Evaluated
HET-CAM Q-Score, with cut-off at 2	0.476-0.701	0.570-0.780
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated
<i>Solids (7-17 depending on endpoint)</i>		
HET-CAM Q-Score	0.578-0.808	0.694-0.875
HET-CAM S-Score	0.060-0.332	-0.009-0.326
HET-CAM Q-Score, with cut-off at 2	0.458-0.694	0.512-0.816
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated
<i>Solutions (14 depending on endpoint)</i>		
HET-CAM Q-Score	0.125-0.678	0.268-0.658
HET-CAM S-Score	Not Evaluated	Not Evaluated
HET-CAM Q-Score, with cut-off at 2	0.238-0.483	0.292-0.493
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated
<i>Liquids (26)</i>		
HET-CAM Q-Score	0.328-0.481	0.489-0.616
HET-CAM S-Score	Not Evaluated	Not Evaluated
HET-CAM Q-Score, with cut-off at 2	0.502-0.550	0.546-0.625
HET-CAM S-Score, with cut-off at 2	Not Evaluated	Not Evaluated

9.1.4 Brantner et al. (2002)

The investigators evaluated compounds and plant extracts for anti-inflammatory properties using the HET-CAM test method. Eggs were initially incubated three days. Then a small hole was drilled into the eggshell, 10 mL of the egg white was removed, and then the hole was sealed. On the opposite side of the egg, the shell was opened with forceps and then covered with parafilm. The egg was then re-incubated for another three days. At that time,

the eggs were treated with SDS to induce an irritant response on the CAM. The investigators evaluated the anti-inflammatory properties of eight steroidal and non-steroidal substances. The ability of the test substances to reduce inflammation was calculated by determining the percent reduction in SDS-induced inflammation of the treated samples. *In vivo* inflammatory and anti-inflammatory responses were determined using the Croton oil test. The investigators indicate that the HET-CAM test method was more sensitive than the *in vivo* test method in determining anti-inflammatory activity of the test substances. However, it is noted that the *in vivo* test method was able to provide dose-response correlations in the substances evaluated, while the HET-CAM test method could not provide clear correlations.

The data from this study was not used in an analysis of HET-CAM test method accuracy, because the response being evaluated was not irritation potential, but anti-inflammatory responses.

9.1.5 Brantom et al. (1997) and Steiling et al. (1999)

Under the auspices of the European Cosmetic, Toiletry, and Perfumery Association (COLIPA), a validation study on alternatives to the *in vivo* ocular toxicity test method was conducted. Using 23 substances that represented cosmetic ingredients (selected from the ECETOC database; ECETOC 1992) and 32 finished products, the validation status of several alternative test methods were evaluated. In this study, the *in vitro* HET-CAM scores (calculated as Q-Score or S-Score) were separated into four different irritancy classifications (slightly, moderately, irritating, and severely). *In vivo* rabbit eye scores were segregated into the same four irritant classes. Comparative *in vivo* rabbit eye test results were obtained from historical sources or concurrent studies conducted in accordance with OECD TG 405. MMAS values were provided for a subset of the tested substances.

Accuracy and interlaboratory reproducibility between *in vivo* classification and *in vitro* classification was determined by the statistic κ^1 . The study indicated that the HET-CAM test method classifications did not accurately predict the *in vivo* classification categories (κ values from 0.268 to 0.541 and κ_Q values from 0.428 to 0.731). The interlaboratory reproducibility (for four laboratories) ranged from 0.342 to 0.607. Analysis indicated that the interlaboratory reproducibility of the test method appeared to be moderately good at the extreme ranges of irritancy (Q-Score of less than 0.8 or greater than 2.0) but was a poor predictor of irritancy of substances with a Q-Score in the middle range (between 0.8 and 2.0).

The data could not be used in the accuracy analysis because individual sample or mean sample *in vitro* scores were not provided in the report. Thus, the test substances could not be classified according to the classification system described in **Section 5.0** and were not used in the accuracy analysis described in **Section 6.0**.

¹ The statistic κ value can either weight all factors equally or use different weightings. For the analysis, three versions of the κ statistic were used: (1) equal weighting for all factors (κ), (2) linear weighting where greater weight was given to the effect of disagreements of more than two classification categories (κ_L), and (3) quadratic weighting where very high weighting was given to the effect of disagreements of more than two classification categories (κ_Q).

9.1.6 Budai et al. (1997) and Budai and Várnagy (2000)

Comparative screening of six pesticides at three concentrations (1%, 10%, and 100%) was conducted to assess the usefulness of the HET-CAM test method when compared to the *in vivo* rabbit eye test method. The *in vitro* scores (calculated as IS(A)) were separated into four different irritancy classifications (no, weak, moderate, severe). Comparative *in vivo* rabbit eye test results were obtained from concurrently run studies conducted in accordance with OECD TG 405. The *in vivo* scores (reported as MAS) were separated into four different irritancy classifications (no or slight, moderate, severe, super). The rationale for the *in vivo* categories was not provided.

The reports indicate that the HET-CAM results showed good correlation to the *in vivo* results. Of the four test substances tested *in vitro* and *in vivo*, three substances were classified in similar categories by both test methods. One test substance was overclassified by the HET-CAM test method.

The data from this study could not be used in the accuracy analysis because individual sample or mean sample *in vitro* scores were not provided in the report. Therefore, the tested substances could not be classified according to the classification system described in **Section 5.0**.

9.1.7 CEC (1991)

A collaborative study on alternative methods to the *in vivo* rabbit eye test was commissioned by the Division Control of Chemicals, Industrial Risks and Biotechnologies of Directorate General Environment, Nuclear Safety, Civil Protection and the Health and Safety Directorate of Directorate General Employment Industrial Relations and Social Affairs. *In vitro* IS values were calculated according to the method of Kalweit et al. (1987) (IS[B] analysis method). A score of greater than 9 was defined as a severe irritant. *In vivo* data were classified according to the EU classification system based on chemical profiles developed for the evaluation.

The authors indicate that the HET-CAM test method performed well in identification of severe irritants (R41 classified substances). However, nonirritants were overclassified. The authors suggest that an improved evaluation may be obtained if dilutions of the test substances were evaluated.

A subset of the data present in this study was used in the BRD. For the accuracy and reliability evaluations described in **Section 6.0** and **Section 7.0**, substances where no *in vivo* rabbit studies were used in the irritancy classification were excluded from consideration.

9.1.8 Dannhardt et al. (1996)

The investigators evaluated whether the HET-CAM test method could be used as a screen for nonsteroidal anti-inflammatory agents. A unique test method protocol was used in which the eggs were incubated for nine days. A small hole was then drilled into the eggshell and the test substance was placed on the CAM using a syringe. The hole was sealed with cement. The eggs were then incubated for 2, 4, or 6 hours. After the incubation period, the eggshell was opened and SDS was placed on the CAM. The time of the start of the irritation response

was then noted. The time of the start of the irritation response of eggs treated with test substance was compared to those treated with negative controls and the relative delay of onset, if any, was determined. The investigators indicate that the method allows the classification of the substances according to their potency; however, correlation with the results of *in vivo* studies was limited.

The data from this study was not be used in an analysis of HET-CAM test method accuracy, because the response being evaluated was not irritation potential, but anti-inflammatory responses.

9.1.9 Demirci et al. (2003)

The investigators evaluated substances isolated from essential oils from aerial parts of *P. linearis* for antiangiogenic and anti-inflammatory properties. The test method used in this evaluation comprised forming an agarose pellet with the test substance and applying the pellet to the CAM surface. The severity of the observed effect was scored on a scale from 0.5 to 1. No comparative *in vivo* studies were conducted in this evaluation. The evaluation showed that application of the agarose pellets containing the test substance was not toxic and did not produce irritant effects.

The data from this study was not be used in an analysis of HET-CAM test method accuracy, because the response being evaluated was not irritation potential, but anti-inflammatory responses.

9.1.10 Demirci et al. (2004)

The investigators evaluated substances isolated from the essential oils from aerial parts of *Origanum onites* L for antiangiogenic and anti-inflammatory properties of the isolated substances. The test method used in this evaluation comprised forming an agarose pellet with the test substance and applying the pellet to the CAM surface. The severity of the observed effect was scored on a scale from 0.5 to 1. No comparative *in vivo* studies were conducted in this evaluation. The evaluation showed that application of the agarose pellets containing the test substance was not toxic and did not produce irritant effects.

The data from this study was not be used in an analysis of HET-CAM test method accuracy because the response being evaluated was not irritation potential, but anti-inflammatory responses.

9.1.11 de Silva et al. (1992)

The investigators evaluated 60 chemicals and 41 cosmetic formulations; the chemicals and components of the formulations tested were not provided in the report. The commercial products classes of the formulations were oils, make-up removal, emulsions, gels, shampoos, and creams and body milk. The chemicals tested were evaluated at 1% and 10% concentrations, while the formulations were tested neat. Of the 41 formulations tested, 20 were rinsed off the CAM 20-seconds after application because they were opaque or colored. In this study, *in vitro* scores were calculated as IS(A) values and classified as described in

Luepke (1985). The comparative *in vivo* rabbit eye scores (calculated as MAS and classified per the EEC classification scheme²) were obtained from published results.

The studies showed that interlaboratory reproducibility was high for test chemicals evaluated at 1% and 10% concentrations. The Spearman's coefficient for both concentrations was greater than 0.9. The results from the *in vitro* analysis were plotted against the EEC categories (tabular data were not provided) and relationship between the two was determined using the Jonckheere-Terpstra test, followed by calculation of the Spearman's coefficient. This analysis yielded a moderate coefficient of 0.726 ($p < 0.0001$). For these substances, HET-CAM had an accuracy of 90%, a sensitivity of 91%, and a specificity of 88%³.

The rank correlation between those formulations that were rinsed and those that were not rinsed were compared to determine the effect of protocol differences. The results obtained were compared to the MAS. The Spearman's coefficients were 0.77 for the non-rinsed formulations and 0.76 for the rinsed formulations.

The *in vitro* data from this study could not be used in the HET-CAM accuracy analysis, because sufficient information on the test substances and *in vitro* score were not provided in the report. The lack of information on the test substances did not allow for categorization of the substances into the irritancy categories defined by the GHS (UN 2003) or EPA (1996).

9.1.12 Djabari et al. (2002)

Investigators combined the HET-CAM test method with histological evaluation of the CAM in an attempt to increase the sensitivity of the test method. The test method was conducted and scored as described in Luepke (1985) (IS[A] analysis method). Immediately after CAM scoring was completed, the central part of the CAM was removed, fixed, and stained with trypan blue to evaluate the state of the blood vessels. Twenty water-soluble test substances (identified as active ingredients in cosmetics) were evaluated undiluted and at a 10% dilution. No comparative *in vivo* studies were conducted in this evaluation.

The report indicates that when the diluted forms of the test substances were evaluated there was no discrepancy between the results of the HET-CAM evaluation and histological evaluation of the CAM. At a 10% concentration, all the substances were classified as nonirritant by the HET-CAM method and no morphological changes were observed by histological or trypan-blue evaluation. When the substances were evaluated undiluted, seven of the substances displayed discrepancies between the results of the HET-CAM evaluation and the histological evaluation. For six of the substances, the HET-CAM evaluation indicated that the substances were nonirritants while the histological evaluation indicated that the substances produced irritation. In the last case, the histological evaluation indicated that the substance produced slight hemorrhages while the visual inspection of the CAM indicated the development of hyperemia. The investigators concluded that inclusion of histological

² No citation is provided in the study regarding the specific guideline used in classifying substances. However, the study indicates that the classifications used in the analysis (Class I, II, and III) correlate to nonirritant, R36, and R41, respectively (EU 1992, 2001).

³ Numbers used to calculate these percentages are only provided in graphical form and no attempt was made to count the points.

examination of the CAM after conducting the HET-CAM assay could increase the sensitivity of the method and provide greater information about the effects produced by the test substance.

The data from this study could not be used in a HET-CAM accuracy analysis (**Section 6.0**) because comparative, *in vivo* data for the test substances was not provided in the report and such data were not located.

9.1.13 Doucet et al. (1999)

Comparative screening of 40 cosmetic formulations was conducted to assess the usefulness of *in vitro* ocular toxicity test methods, including the HET-CAM test method, when compared to the currently accepted *in vivo* rabbit eye test method. The formulations were classified as skin care products (10), sunscreen products (10), surfactant based products (10), and alcoholic products (10). In this study, the *in vitro* scores were calculated as IS(A) values; value greater than five was defined as an irritant. Comparative *in vivo* results (calculated as MMAS) were calculated from concurrently conducted studies run according to the method described by Draize et al. (1944). A substance with an MMAS value greater than 15 was defined as an irritant. There was no rationale provided for the classification and cut-off values used.

Correlation between the HET-CAM IS(A) values and MMAS values yielded a κ value of 0.58. The linear correlation between these values was statistically significant ($p < 0.001$) and Pearson's coefficient was 0.72. The calculated residual standard deviation, however, was large. Evaluation of accuracy parameters yielded the following values: accuracy: 80%, sensitivity: 100%, specificity: 56%, false positive rate: 44%, false negative rate: 0%. Of the substances that were identified as false positives, four were skin care products and four were sunscreen products.

The data from this study could not be used in an analysis of HET-CAM test method accuracy because the *in vivo* data provided in the report was insufficient to classify the substances according to the GHS (UN 2003), EPA (1996), or EU (2001) classification systems.

9.1.14 Gettings et al. (1991, 1994, 1996) and Lordo et al. (1999)

The CTFA developed an Evaluation of Alternatives Program, with the intent to provide industry with sufficient information on the performance of a series of potential alternatives to the *in vivo* ocular toxicity test method. This effort was a multi-year, multi-phase effort, with different product-types tested in each phase. The evaluation focused on assessing the accuracy of alternative test methods when compared to the FHSA classification system (CPSC 1988).

The initial phase evaluated a set of ten generic hydroalcoholic formulations (Gettings et al. 1991). In this phase, *in vitro* IS values were calculated via two mathematical methods (Bartnik et al. 1987; Kalweit et al. 1987). A substance with an IS value greater than 300 or 10, respectively, was defined as an irritant. The *in vivo* results were expressed as irritants or nonirritants, based on the FHSA regulatory classification system. No *in vivo* scores (e.g., MAS, Draize scores, animal scores) were provided in the report. In this phase of the

evaluation, the HET-CAM test method displayed 100% (5/5) sensitivity and 100% (4/4) specificity when compared to the *in vivo* classification based on the FHSA regulatory classification system (CPSC 1988).

The data from this report were re-evaluated since the *in vivo* data was classified according to the FHSA classification system. Based on additional data obtained from CTFA and the FDA, the ability of the HET-CAM test method to accurately identify ocular corrosives and severe irritants, as defined by the GHS (UN 2003), EPA (1996) and EU (2001) classification systems, is described in **Section 6.0**.

The second phase of the evaluation focused on a set of 18 generic oil-water emulsion formulations (Gettings et al. 1994). In this study, *in vitro* IS(A) and IS(B) values were calculated. In this evaluation, a substance with an IS(A) value equal to or greater than 4.8 or an IS(B) value equal to or greater than 5 was defined as an irritant. As in the previous phase, test substances were classified as either irritants or nonirritants according to the FHSA classification system. In this phase, when the *in vitro* data were transformed using the IS(B) analysis method, the sensitivity was 100% (5/5) and the specificity was 85% (11/13). When the *in vitro* results were transformed using the IS(A) analysis method, the sensitivity was 80% (4/5) and the specificity was 77% (10/13).

The data from this report was re-evaluated since the *in vivo* data was classified according to the FHSA classification system. Based on additional data obtained from CTFA, the ability of the HET-CAM test method to accurately identify ocular corrosives and severe irritants, as defined by the GHS (UN 2003), EPA (1996) and EU (2001) classification systems, is described in **Section 6.0**.

The third phase of the evaluation focused on a set of 25 generic surfactant-based formulations (Gettings et al. 1996). In this study, *in vitro* IS(A) and IS(B) values were calculated. A substance with an IS value equal to or greater than 5.1 or 4.83, respectively, was defined as an irritant. A ratio of IS to ITC also was evaluated. Substances with an IS/ITC value of equal to or greater than 3.0 was defined as an irritant.

In this evaluation, the formulations were classified as irritants or nonirritants based on each of the models described. Accuracy assessments were then conducted for each model. Using the IS(B) analysis method, the sensitivity was 94% (17/18) and the specificity was 71% (5/7). Using the IS(A) analysis method, the sensitivity was 94% (17/18) and the specificity was 100% (7/7). Using the IS/ITC ratio model, the sensitivity was 100% (18/18) and the specificity was 71% (5/7).

Since the *in vivo* data was classified according to the FHSA classification system, all of the data from this report was re-evaluated. Based on additional data obtained from CTFA, the ability of the HET-CAM test method to accurately identify ocular corrosives and severe irritants, as defined by the GHS (UN 2003), EPA (1996) and EU (2001) classification systems, is described in **Section 6.0**.

In the report by Lordo et al. (1999), the investigators evaluated the precision and extent of random variations associated with the regression fits determined with the data described in the Gettings et al. reports. The sources of variation around each of the regression models were evaluated by estimating the components of total variation associated with predicting MAS for each phase of the CTFA evaluation.

From the evaluation, the greatest source of variability associated with predicting the MAS was due to random variations around the prediction models (70% to 90%) for each of the phases. Generally, variability between *in vitro* replicates and variability between MAS replicates contributed only a minor proportion to the total variability associated with the models for the test substances. The authors conclude that the contribution of the latter two variability components could be decreased by increasing the number of replicates performed for each test formulation. However, it would have little impact on the overall precision of the prediction models developed by Gettings et al. (1996).

9.1.15 Gilleron et al. (1996)

This report discusses an alternative test method protocol for the HET-CAM test method. In this method, the investigator used a TSA to confine the test substance to a section of the CAM. The report discusses the evaluation of 46 substances. The *in vitro* scores were calculated as IS(B) scores. A substance with an IS(B) value equal to or greater than 5.0 was defined as an irritant. The *in vivo* results were reported as MAS. Additionally, the irritancy potential of each test substance was classified based on the EU classification system (EU 1992).

The correlation between IS(B) and MAS values was moderate and statistically significant ($r = 0.58$, $p \leq 0.001$). The best correlation was obtained between the total IS(B) value and the *in vivo* conjunctival score ($r = 0.68$, $p \leq 0.001$). Correlation coefficients between *in vivo* and *in vitro* results, based on physical properties of the test substances, also were conducted ($r = 0.72$ for solids; $r = 0.78$ for liquids; and $r = 0.93$ for surfactants).

Accuracy analysis with the test substances indicated that the HET-CAM method (with the use of TSA) exhibited high sensitivity (92.3% [12/13]) but low specificity (54.5% [18/33]) in classifying substances as irritants or nonirritants. The results of an assessment of the accuracy of the test method for solids, liquids, and surfactants are provided in **Table 9-2**.

Table 9-2 Accuracy Statistics for Test Substances Evaluated in Gilleron et al. (1996)

Statistic	Solids	Liquids	Surfactants
Accuracy	88% (15/17) ¹	38% (8/21)	88% (7/8)
Specificity	92% (11/12)	24% (4/17)	75% (3/4)
Sensitivity	80% (4/5)	100% (4/4)	100% (4/4)
False Negative	20% (1/5)	0% (0/4)	0% (0/4)
False Positive	8% (1/12)	76% (13/17)	25% (1/4)

¹Numbers in parentheses were used to calculate the percentages.

Data obtained from the report were reclassified based on chemical class and properties of interest. The average HET-CAM IS(B) values and EU irritancy classification provided in the report were used in the analyses described in **Section 6.0**. In response to a request from

NICEATM, Drs. Vanparys and Goethem forwarded raw *in vitro* data that was used for the reliability analysis described in **Section 7.0**.

9.1.16 Gilleron et al. (1997)

This report describes a follow up evaluation to the Gilleron et al. (1996) study. In this study, the investigators evaluated 60 substances. The substances included 28 liquids, 20 solids, and 12 surfactants. *In vitro* values were calculated as IS(B) scores, and a substance with an IS value equal to or greater than 5.0 was defined as an irritant. The *in vivo* scores (calculated as MMAS) were calculated from published data. A substance with a MMAS equal to or greater than 15.0 was defined as an irritant.

The total IS(B) and individual HET-CAM endpoints were compared to the total MMAS value. Correlation analyses indicated that no good correlation was observed. The relationship between MMAS and various physicochemical properties (e.g., solids, liquids, surfactants) also was low ($r = 0.29$ to 0.38)

An accuracy analysis of the data indicated that the HET-CAM method (with the use of TSA) exhibited moderate accuracy (80% [48/60]), high sensitivity (96% [45/47]), and low specificity (23% [3/13]). The results of an assessment of the accuracy of the test method for solids, liquids, and surfactants are provided in **Table 9-3**.

Table 9-3 Accuracy Statistics for Test Substances Evaluated in Gilleron et al. (1997)

Statistic	Solids	Liquids	Surfactants
Accuracy	90% (18/20) ^a	75% (21/28)	75% (9/12)
Specificity	100% (3/3)	0% (0/7)	0% (0/3)
Sensitivity	88% (15/17)	100% (21/21)	100% (9/9)
False Negative	12% (2/17)	0% (0/21)	0% (0/9)
False Positive	0% (0/3)	100% (7/7)	100% (3/3)

^a Numbers in parentheses were used to calculate the percentages.

Data obtained from the report were reclassified based on chemical class and properties of interest. *In vivo* data for tested substances were obtained from ECETOC (1998). These *in vitro* and *in vivo* data were then used in the analyses described in **Section 6.0**. In response to a request from NICEATM, Drs. Vanparys and Goethem forwarded raw *in vitro* data that was used for the reliability analysis described in **Section 7.0**.

9.1.17 Hagino et al. (1991)

Investigators conducted a comparative screening of 12 surfactants (evaluated as 10% aqueous solutions) to assess the usefulness of the HET-CAM test method, when compared to the *in vivo* rabbit eye test method. The surfactants were classified as cationic (3), anionic (5), nonionic (2), and amphoteric (2). In this study, *in vitro* scores were calculated as IS(A) values. The *in vivo* rabbit eye study scores (presented as the maximum total Draize score) were calculated from concurrently run studies conducted according to the method described by Draize et al. (1944). The results from this study indicated that there was good correlation between the IS(A) value and the maximum total Draize score ($r = 0.86$).

The data in the report was presented in graphical form and no attempt was made to extrapolate the points to estimate mean HET-CAM IS(A) values. Since the *in vitro* scores were not provided, the tested substances could not be classified according to the classification system described in **Section 5.0**.

9.1.18 Hagino et al. (1993)

In this evaluation, the investigators compared the HET-CAM results of 12 substances to the *in vivo* rabbit eye test method. The 12 substances comprised a variety of physical forms (liquids, powders, and emulsions) and solubilities (seven of the 12 substances were not soluble in water). All but two substances were tested undiluted. No rationale was provided in the report as to the selection of the test substances, the number of substances tested, or the concentration tested. *In vitro* scores were calculated as IS(A) values. The *in vivo* scores (presented as the MAS) were obtained from published studies that used techniques that were similar to the method described by Draize et al. (1944).

There was good correlation between the IS(A) value and the maximum total Draize score ($r = 0.90$). Increasing concentrations of a test substance (ethanol) were shown to produce increased response in HET-CAM, suggesting that the method could assess dose response relationships. The data for this evaluation were combined with the data from a previous evaluation (Hagino et al. 1991) and then separated by solubility (water soluble and non-water soluble). The responses in the HET-CAM test method for these two classes were relatively similar. Overall, the correlation coefficient for all test substances was 0.80.

The data in the report was presented in graphical form and no attempt was made to extrapolate the points to estimate mean HET-CAM IS(A) values. Since the *in vitro* scores were not provided, the tested substances could not be classified according to the classification system described in **Section 5.0**.

9.1.19 Hagino et al. (1999) and Ohno et al. (1999)

Two types of CAM assays, HET-CAM and chorioallantoic membrane-trypan blue staining (CAM-TB), were evaluated by investigators as alternative methods to the *in vivo* rabbit eye test method. The validation effort was composed of three phases where a total of 39 test substances were evaluated. The test methods were evaluated in five different laboratories. In this study, *in vitro* scores were calculated as IS(A) values; value equal to or greater than 7.0 was defined as an irritant. The *in vivo* scores (calculated as the maximum total Draize score) were calculated from published studies that were conducted according to the method described by Draize et al. (1944). A substance with an MAS value greater than 15 was defined as an irritant. According to investigators, the *in vitro* cut-off was set arbitrarily based on the distribution pattern of the substances while the *in vivo* cut-off was set according to the classification system defined by Kay and Calandra (1962).

The results showed that HET-CAM correctly identified the irritancy potential of 46 of the 52 test substances. Five chemicals were classified as false positives and one chemical was classified as a false negative. Correlation analysis indicated that the rank correlation coefficient between the HET-CAM IS(A) values and MAS was 0.802. Spearman's rank correlation coefficient also was high for the relationship between the IS(A) value obtained

for each testing laboratory when compared the mean IS(A) value of all of the testing laboratories (0.856 to 0.950).

The data from this report was re-evaluated since the classification system described in the report did not have a severe irritant classification, which is the focus of this evaluation. In response to a request from NICEATM, Dr. Yasuo Ohno of NIHS forwarded raw *in vivo* data. Using this data, the ability of the HET-CAM test method to accurately identify severe irritants was provided in **Section 6.0** and reproducibility results were provided in **Section 7.0**.

9.1.20 Kalweit et al. (1987)

This report describes the results from the first preliminary trial of a national validation project to validate alternative methods to the *in vivo* ocular toxicity test method. During this phase, the HET-CAM test method was established in participating laboratories. In this evaluation, two substances (SDS and triethanolamine) were evaluated in six different laboratories. *In vitro* scores were calculated using the IS(B) analysis method. *In vivo* studies were not conducted for the analysis described in the report.

The report stated that there was close agreement of the results with a high concentration of SDS (1%). Five of the six laboratories classified the test substance as a strong irritant. At lower concentrations of SDS (0.1% to 0.5%), a clear classification of the irritancy potential of the test substance was not possible. Investigators stated that similar results were observed with triethanolamine.

The data from this study could not be used in the accuracy analysis because individual sample or mean IS(B) values were not provided in the report. Therefore, the tested substances could not be classified according to the system described in **Section 5.0**.

9.1.21 Kalweit et al. (1990)

This second report describes additional results from the preliminary phase of a national validation project to validate alternative methods to the *in vivo* ocular toxicity test method. During this phase, two HET-CAM test trials were conducted to test the protocols and software developed for the evaluation. Five substances were evaluated and interlaboratory reproducibility was determined. The substances tested were zinc pyridinethione, 2-butoxyethanol, dimethylsulfoxide, triethanolamine, and SDS. In this study, the *in vitro* scores were calculated as IS(B). *In vivo* studies were not conducted and *in vivo* data was not used in the analysis. The report stated that there were considerable differences between results obtained by trained and less experienced investigators. No additional statistical analyses or evaluations were provided in the report.

The data from this study could not be used in the accuracy analysis because individual sample or mean sample scores were not provided. Since the *in vitro* scores were not provided, the tested substances could not be classified according to the classification system described in **Section 5.0**.

9.1.22 Kojima et al. (1995)

Investigators evaluated seven alternative test methods to the *in vivo* ocular toxicity test method. Twenty-four test substances were evaluated; six substances were cationic surfactants, five substances were anionic surfactants, seven substances were nonionic surfactants, two substances were amphoteric surfactants, and four substances were solvents. *In vitro* HET-CAM test method data were calculated as IS(A) values. Concurrent *in vivo* studies were conducted similar to what was previously described in Draize et al. (1944). In this evaluation, three female rabbits were observed at 1, 3, 6, 24, 48, 72, 96, and 168 hours post application of the test substance. The maximal Draize rabbit eye irritation score (MDES; calculated in a manner similar to the MAS and MMAS) was then calculated. The investigators concluded that there was a moderate relationship between the HET-CAM IS(A) values and MDES (correlation coefficient = 0.824).

The analysis described in this report could not be used directly in an analysis of HET-CAM accuracy because the *in vivo* data was insufficient to classify the substances according to one of the three ocular irritation classification systems used in this analysis. However some of the data from this study was re-analyzed, using historical *in vivo* data from other sources. The results of this re-analysis are provided in **Section 6.0**.

9.1.23 Lawrence et al. (1990)

Investigators conducted comparative screening of 34 substances to assess the usefulness of the HET-CAM test method. The substances ranged from single chemicals to fully formulated products (e.g., shampoos and industrial detergent cleaners). Results from the HET-CAM test method were expressed as *in vitro* irritation classification categories described in Luepke (1985). The *in vivo* results were classified into eight irritancy categories. The data used for the *in vivo* classifications were based on published data from studies that were conducted according to the method described by Draize et al. (1944). The investigators reported that there was not a good correlation between the *in vitro* and *in vivo* results.

The data from this study could not be used in the HET-CAM accuracy analysis, because the *in vivo* data was insufficient to classify the substances according to one of the three classification systems reviewed. The test substances evaluated also were not identified; therefore the use of historical *in vivo* rabbit eye data to conduct an accuracy analysis was not possible.

9.1.24 Lönnroth et al. (1999)

The irritation potentials of eight dental polymer products were tested using the HET-CAM test method. *In vitro* data were evaluated using the IS(B) analysis method. The report did not evaluate the *in vivo* effects of these test substances or correlate *in vitro* results with *in vivo* results. The results showed that the liquid components of all the products had strong irritation potential but the powder suspensions and extracts had no effect.

The data from this study could not be used in the HET-CAM accuracy analysis because *in vivo* data and ocular irritancy classification information, as defined by the GHS (UN 2003), EPA (1996), and EU (2001) classification systems, were not provided.

9.1.25 Luepke (1985)

This report provides the initial description of the HET-CAM test method. Chemicals and formulations (vehicles, antimicrobial agents, oxidation dyes, and commercial shampoos) were tested. *In vitro* irritancy classifications of tested substance, not IS values, were provided in the report. The *in vivo* irritancy classifications consisted of four categories; however, information on how the *in vivo* data was collected was not provided.

The author concluded that the HET-CAM test method was capable of demonstrating the mucous membrane irritating potencies of substances. The investigator indicates that the method was useful for screening large numbers of compounds.

The data from this study could not be used in the HET-CAM accuracy analysis because *in vivo* data for the test substances were not available and historical *in vivo* rabbit eye data could not be located.

9.1.26 Luepke and Kemper (1986)

In this study, the investigators evaluated the usefulness of the HET-CAM test method using about 190 substances and formulations. The investigators noted that there was good correlation between the *in vitro* and *in vivo* data and that there was a high level of reproducibility between laboratories.

The data from this study could not be used in the HET-CAM accuracy analysis because *in vivo* rabbit eye data for the test substances were not available. Additionally, the identities of the substances tested were not provided and, therefore, historical *in vivo* rabbit eye data could not be used to conduct an accuracy analysis.

9.1.27 Macián et al (1996)

The investigators report evaluated the toxic effects of a group of synthetic polyoxyethylene nonionic surfactants, which were developed by the investigators. Ocular toxicity potential was evaluated with the HET-CAM test method. In this study, the *in vitro* scores (reported as IS(B)) were calculated using a formula that evaluated the irritancy potential index. *In vivo* rabbit eye studies were not conducted and *in vivo* rabbit eye data were not used in the analysis. The report stated that the test substances were weak to moderate irritants based on the results from the HET-CAM test method.

Since the chemicals evaluated in this study were novel, historical data for the effects of these substances in ocular irritation tests could not be obtained. The lack of comparative *in vivo* data precluded the use of this study in the HET-CAM accuracy analysis.

9.1.28 Reinhardt et al. (1987)

The investigators conducted a comparative screening of 24 surfactants to assess the usefulness of the HET-CAM test method, when compared to the *in vivo* guinea pig eye test method. The selected surfactants all induced a similar range of *in vivo* eye irritation, which was defined as slight. All test materials were tested as a concentration of 300 mM or 10% mixtures. In this study, the *in vitro* scores were calculated as IS(A) values and classified as described Luepke (1985). The *in vivo* scores (reported as the maximum total Draize score

over a 24 hour period) were calculated in concurrent studies on guinea pigs. Eye irritation scoring was analogous to the rabbit *in vivo* eye ocular toxicity test method (Draize et al. 1944) and irritation severity was assessed according to Kay and Calandra (1962). The results showed that the HET-CAM test method was poor in predicting the eye irritation potential of anionic surfactants; the method overpredicted the severity of irritation produced by these test substances.

The data from this study could not be used in the HET-CAM accuracy analysis because comparative *in vivo* data were not provided in the report.

9.1.29 Rougier et al. (1992)

The investigators conducted comparative screening of 41 surfactants and surfactant based formulations to assess the usefulness of several *in vitro* ocular toxicity test methods, including HET-CAM. *In vitro* results were calculated as IS(A) values. The *in vivo* scores (reported as MAS) were based on published data. Spearman Rank correlations were calculated across various data sets and *in vitro* and *in vivo* endpoints.

The analyses showed a high degree of correlation between the HET-CAM hemorrhage score and the MAS value for surfactants and surfactant-based formulations ($r_s = 0.98$ and 0.95 , respectively). The overall rank correlation coefficient for all 41 substances was 0.96 .

The data from this study could not be used in the HET-CAM accuracy analysis because the *in vivo* data was insufficient to classify the substances according to the GHS (UN 2003), EPA (1996), or EU (2001) classification system, and historical *in vivo* rabbit eye data for the substances tested could not be located.

9.1.30 Schlage et al. (1999)

The investigators evaluated the use of the HET-CAM test method to determine the irritant potential of cigarette mainstream and sidestream smoke. In this study, *in vitro* IS values were calculated and classified as described in Kalweit (1985). *In vivo* results were not evaluated for this analysis. The data from this study could not be used in the HET-CAM accuracy analysis because historical *in vivo* data were not located.

9.1.31 Spielmann et al. (1991, 1993, 1996)

Starting in 1988, a national validation study on two alternative ocular toxicity test methods was initiated by ZEBET. Spielmann et al. (1991) described the interlaboratory assessment and the database development. In this phase, 32 coded substances that represented a variety of chemical and toxicological properties were evaluated in 12 laboratories to assess interlaboratory reproducibility of HET-CAM. All but four test substances were evaluated at 10% concentrations; the four remaining substances were evaluated at concentrations ranging from 0.5% to 100%. Additionally, the lowest concentration required to produce a slight reaction on the CAM was determined. These studies were conducted in two laboratories with experience in the test method. In this study, the *in vitro* scores were calculated as IS(B) values. The irritation classification scheme used in the evaluation was performed according to Luepke (1985). The *in vivo* results, expressed as irritation classification categories (e.g.,

slight, moderate, severe) developed by the investigators, were obtained from published studies that were conducted as described by Draize et al. (1944).

The results indicated that of the 27 test substances that were evaluated by the HET-CAM test method, 16 (59%) were classified correctly (nine positives and seven negative) by 75% of the testing laboratories. There were five false positives and one false negative result. In addition, five of the test substances evaluated did not yield the same classification in at least 75% of the testing laboratories and thus could not be classified.

In the second phase of the national validation study conducted by ZEBET, 136 coded substances that represented a variety of chemical and toxicological properties were discussed (Spielmann et al. 1993). The substances tested were evaluated at 10% concentration. Additionally, the lowest concentration required to produce a slight reaction on the CAM also was determined. The studies were conducted in two laboratories (of seven possible laboratories) with experience using the test method. In this study the *in vitro* scores were calculated as IS(B) values. The mean value of three eggs was used for each test substance. The irritation classification scheme used in the evaluation used both the IS and ITC values. The *in vivo* scores (classified per the EU classification scheme [EU 1992]) were obtained from published studies that were conducted as described by Draize et al. (1944) in compliance with GLP guidelines.

Of the 136 substances tested, 46 were classified as severe irritants (R41) based on *in vivo* studies. Of these 46 substances, both test laboratories correctly identified 22 substances as R41. A majority of the remaining substances (15) were classified as nonirritant or moderately irritant by both test laboratories. Correct identification of the nonirritants was 80%, while identification of R36 labeled chemicals was 10%. The authors indicate that the HET-CAM test method could be incorporated into the OCEG TG 405 testing scheme and be used to reduce the suffering associated with the evaluation of ocular corrosives and severe irritants.

In the third phase of the national validation study conducted by ZEBET, 200 coded substances that represented a variety of chemical and toxicological properties were discussed (Spielmann et al. 1996). The chemicals tested were evaluated undiluted and at a 10% concentration. Additionally, the lowest concentration required to produce a slight reaction on the CAM also was determined. The studies were conducted in two laboratories (of seven possible laboratories) with experience in the test method. The *in vitro* scores were calculated as IS(B) values. The irritation classification scheme used in the evaluation considered both IS and ITC values. The *in vivo* results (expressed as irritation severity categories defined by the EU classification system [EU1992]) were obtained from published results and unpublished results provided by chemical and pharmaceutical companies; the studies were conducted as described by Draize et al. (1944) and in compliance with GLP guidelines.

Of the 200 substances tested, 118 were used in the evaluation of the ability of the HET-CAM test method to identify severe irritants (R41). An assessment of the accuracy statistics of the

test method, based on these 118 chemicals, was conducted by the authors who reported a sensitivity of 41% (19/45), specificity of 89% (65/73), and false positive rate of 7%⁴.

Additional endpoints were derived from the calculated scores to conduct discriminant analyses. These analyses showed that use of the mtc10 endpoint (mean detection time for appearance of coagulation when using a 10% solution) correlated better with severe irritants than any other evaluated endpoints (sensitivity: 52.1% [25/48], specificity: 88.3% [83/84], false positive rate: 7.8%⁵). The power of this endpoint to discriminate between R41 and non-R41 chemicals was 10 times higher than that of the next best endpoint (mtc100 [mean detection time for appearance of coagulation when using 100%]). The authors note that the mtc10 endpoint was better suited to identifying R41 irritants than the original prediction model (using IS and ITC values).

The authors additionally proposed several sequential testing strategies to classify ocular corrosives and severe irritants. According to the authors, the best discrimination of R41 substances occurred when the solubility of the substance was determined in water and/or oil. Based on the level of solubility, one of three different procedures could be followed. The three procedures (described in the report) combined endpoints from the HET-CAM test method (mtc10 and/or mtc100) with endpoints of the Neutral Red Uptake test method. Additional details regarding the procedures are provided in the report.

Based on a request from NICEATM, *in vivo* and *in vitro* data were obtained from the authors. The obtained data were re-evaluated using the classification rules described in **Section 4.0** and **Section 5.0**. These data were then used in the analyses described in **Section 6.0** and **Section 7.0**.

9.1.32 Spielmann et al. (1997)

This report describes a retrospective study of the HET-CAM test method that was conducted by the U.S. Interagency Regulatory Alternatives Group (IRAG). In response to a request by IRAG to the scientific community, five sets of data using three different test method protocols were submitted. The substances represented a broad spectrum of industrial chemicals. Overall information about the solubility of the test substances, the pH ranges, chemical classes, and physical form were provided for each set of submitted data. Individual *in vitro* scores were not provided in the report. The *in vivo/in vitro* correlation between HET-CAM scores and five *in vivo* endpoints (cornea/opacity, iris, erythema, chemosis, discharge) were calculated and reported. In this evaluation *in vitro* scores were compared to the non-weighted mean of modified maximum individual score (\sum MMMIS). To assess *in vitro/in vivo* correlations between different *in vitro* endpoints and *in vivo* scores, Pearson's single and Pearson's partial linear regressions were calculated.

Based on Pearson's correlation coefficients, the HET-CAM scores were highly correlated to the \sum MMMIS ($r_p = 0.607-0.913$) for four of the five HET-CAM data sets that were submitted. *In vitro/in vivo* correlations indicated that, overall, corneal opacity and iritis showed better correlation with *in vitro* endpoints than other adverse effects in the eye. When

⁴ With a specificity rate of 89%, the false positive rate would be expected to be 11%.

⁵ With a specificity rate of 88.3%, the false positive rate would be expected to be 10.7%

a single linear regression was used to correlate *in vitro* scores to *in vivo* endpoints, the prediction rates ranged from 16 to 36% for erythema and 84% to 88% for chemosis. However, when a partial regression was used, the prediction rates ranged from 92% to 100%. Additionally, when analyses were restricted to a protocol conducted by a single test laboratory and a limited chemical class, the tissue damage prediction was > 95%. The HET-CAM test method showed the best prediction with surfactants and surfactant-based formulations.

Data from this evaluation was encompassed by other studies that were used in evaluating accuracy and reliability of the HET-CAM test method in **Section 6.0** and **Section 7.0**,

9.1.33 Sterzel et al. (1990)

Comparative screening of 10 substances was conducted to assess the usefulness of the HET-CAM test method. In this study, the *in vitro* scores were calculated as IS(A) values. The *in vivo* rabbit eye scores were obtained from concurrent rabbit studies conducted in accordance with OECD TG 405. The authors concluded that the results indicated that the HET-CAM test method could identify irritating test substances. The study also indicated that the test method was highly sensitive and the authors concluded that, due to this heightened sensitivity, only substances that cause irritation *in vitro* over a 100-fold concentration range should be specified as potential eye irritants.

The data from this study could not be used in the HET-CAM accuracy analysis, because sufficient *in vitro* information was not available in the report. Additionally, sufficient *in vivo* data was not provided to allow for classification of the test substances into at least one of the three classification schemes (GHS [UN 2003], EPA [1996], or EU [2001]), used in this analysis.

9.1.34 van Erp et al. (1990)

The HET-CAM test method was combined with the use of bovine eyes to assess the irritancy potential of chemical substances towards the conjunctivae and the cornea, respectively. The screening method was referred to as BECAM. In this study, *in vitro* scores were calculated as IS(A) values. *In vivo* rabbit eye studies were performed concurrently in accordance with OECD TG 405. The *in vivo* effects were scored according to the Draize scoring system (Draize et al. 1944) and the scores were classified according to the classification scheme of Kay and Calandra (1962). *In vitro* classification of test substances was compared to the EEC classification system (EEC 1983, 1984).

The authors concluded that the combination of HET-CAM and BCOP *in vitro* results showed a good correlation with the *in vivo* classification results. The investigators noted three limitations with the BECAM screening method: (1) inability of the assay to determine effects on the iris, (2) substances that contained a carbamate group or adhere firmly to the bovine cornea or CAM might generate false results, and (3) it was not possible to determine if a severe effect *in vitro* would result in either reversible or irreversible injury of the eye.

The data from this study could not be used in the HET-CAM accuracy analysis because sufficient *in vitro* information was not available in the report. Additionally, sufficient *in vivo*

data was not provided to allow for classification of the test substances into at least one of the three classification schemes used in this analysis.

9.1.35 Vinardell and García (2000)

In this study, the HET-CAM test method was modified to include evaluation of damage to the membrane, which was quantified as the amount of trypan blue adsorbed into the CAM. The modified method was used to assess the potential ocular irritation caused by mixtures of liquid scintillation cocktails. Adsorbed trypan blue was quantified using a spectrophotometer. The test substances were evaluated at concentrations ranging from 12.5% to 100%. The *in vivo* rabbit eye scores (expressed as the Draize score) were obtained from concurrent studies conducted as described by Draize et al. (1944).

The analysis in the report showed that there was good correlation between test substance concentration and the amount of trypan blue adsorbed into the CAM. Additionally, there was good correlation observed between the amount of trypan blue adsorbed and *in vivo* ocular irritation ($r^2 = 0.9722$).

Individual sample or mean sample *in vitro* scores were not provided in the report. Therefore, classification of substances into standardized *in vitro* irritancy classification categories was not possible. Additionally, the *in vivo* rabbit study data was insufficient to classify the substances according to one of the three classification systems evaluated in this analysis.

9.1.36 Vinardell and Macián (1994)

The irritancy potential of substances used as vehicles (six chemicals) and disinfectant solutions (six solutions) were evaluated in the HET-CAM test method to assess ocular irritancy potential. In this study, the *in vitro* scores were calculated as IS(B) values. *In vivo* scores for the six solutions were obtained from concurrent *in vivo* rabbit eye studies that were conducted in accordance with Draize et al. (1944). For the rabbit studies, the ocular irritation index was calculated, which corresponded to the highest total value obtained after a single application. The results of the *in vivo* test were categorized according to Le Moullet et al. (1976) and a previous classification scheme implemented by the EPA (1974).

In this study, four of the six vehicles were classified as nonirritants or weak irritants, while the remaining two vehicles (0.1 N NaOH and 1% SLS) were classified as severe irritants. These results were not compared with *in vivo* rabbit eye test results. The study indicated that four of the six tested disinfectant solutions gave similar results *in vitro* and *in vivo* (when classified by the Le Moullet et al. or EPA [1974] classification systems).

The disinfectant solution test data from this study could not be used in the accuracy analysis because sufficient information on the test formulations was not provided for additional analysis. *In vivo* data for some of the vehicles tested were obtained from published sources (e.g., ECETOC); a HET-CAM accuracy analysis of these substances is provided in **Section 6.0**.

9.1.37 Vives et al. (1997)

The ocular irritation potentials of six anionic and nonionic surfactants, which were derived from lysine, were evaluated in the HET-CAM test method. The focus of this evaluation was to correlate irritation potential with structural characteristics of the surfactants in order to develop a less irritating surfactant. In this study, *in vitro* scores were calculated as IS(B) values. The ocular effects of these substances *in vivo* were not evaluated. This evaluation showed that anionic surfactants showed higher irritation potential than nonionic surfactants. However, the presence of lysine as a counterion reduced the irritancy potential of anionic surfactants.

The data from this study could not be used in the HET-CAM accuracy analysis because effects of these substances in the *in vivo* rabbit test were not evaluated and historical rabbit test method information on these surfactants could not be located.

9.1.38 Wilson and Steck (2000)

A modified HET-CAM test method protocol was used by the investigators to assess the anti-irritant properties of plant extracts. The investigators measured delays in the onset of vascular hemorrhage, membrane lysis, and membrane coagulation relative to the effect of the irritant (15% lactic acid) alone. In this study, *in vitro* scores were calculated as IS(B) values. An anti-irritation score (AIS) then was calculated which represented the time of onset of one of the measured endpoints with pretreatment of a test substance compared to the onset of the measured endpoint without pretreatment of the test substance. The three AIS values were used in describing the anti-irritant potential of the test substances. The *in vivo* results were obtained from studies on human volunteers and the effect was evaluated over a 24-hour period.

The data from this study could not be used in the HET-CAM accuracy analysis because the response being evaluated was not ocular irritation potential, but anti-irritant responses.

9.1.39 Worth and Cronin (2001)

The investigators developed prediction models to explore the possibility of distinguishing between eye irritants (as expressed by the EU classification system [EU 1993]) and nonirritants, by using *in vitro* endpoints of the HET-CAM test method and the neutral red uptake test. The investigators used the *in vitro* data published in the report by Spielmann et al. (1996) to develop the prediction models. The quality of each prediction model was determined by applying it to a training set of 129 chemicals and by expressing the goodness of fit in terms of the sensitivity, specificity, concordance, false negative rate, false positive rate, negative predictivity, and positive predictivity of the prediction models.

Four prediction models were developed by the authors. Using a training set of 129 chemicals, the investigators determined that a combination of three endpoints provided the best prediction of *in vivo* ocular irritation. The prediction model used was:

If $3.63 \log(\text{TH}_{10}) + 2.10 \log(\text{TH}_{10}) + 0.94 \log(\text{IC}_{50}) < 11.87$, predict Irritant;
otherwise predict Nonirritant

where

TH10 = mean detection time for hemorrhage with a 10% solution
IC50 = concentration of test chemical (mg/mL) resulting in 50% inhibition of neutral red uptake in 3T3 cells⁶

Accuracy statistics indicate that this prediction model, using the training set, had an accuracy rate of 81%, specificity of 90%, sensitivity of 69%, false negative rate of 31%, and a false positive rate of 10%⁷.

The HET-CAM data described in this report was initially described in Spielmann et al. (1996); these data are considered in **Section 6.0**.

9.2 Data Received in Response to the ICCVAM *Federal Register* Notice or from Study Authors

NICEATM staff made attempts to obtain original HET-CAM data for substances that also had been tested *in vivo* using the standard rabbit eye test. A *FR* notice (Vol. 69, No. 57, pp. 13589-12861; available at <http://iccvam.niehs.nih.gov/methods/eyeirrit.htm>), requesting original HET-CAM (and comparative *in vivo* rabbit) data was published on March 24, 2004. A second *FR* notice for original HET-CAM (and comparative *in vivo* rabbit) was published on February 28, 2005 (Vol. 69, No. 38, pp. 9661-9662). In addition, NICEATM staff contacted authors of selected published HET-CAM studies to request the original HET-CAM data (Gettings et al. 1991, 1994, 1996; Gilleron et al. 1996, 1997; Spielmann et al. 1996; Hagino et al. 1999). In response to these efforts, the following data were obtained.

In vivo data was submitted by the CTFA for the studies described in Gettings et al. (1991, 1994, 1996). Individual animal responses for the days that the animals were observed were provided. This data was used to identify the ocular irritant potential (based on the GHS [UN 2003], EPA [1998], or EU [2001] classification systems) of the test substances for each formulation evaluated. Using this information, combined with the results provided in the published literature, the accuracy of each version of the HET-CAM test method used in these reports was determined and these results are provided in **Section 6.0**.

In vivo data was submitted by Dr. Yasuo Ohno, of the National Institute of Health Sciences in Japan, for the test substances and test concentrations used in the evaluation described in Hagino et al. (1999). Individual animal responses for the days that the animals were observed were provided. This data was used to identify the ocular irritant potential (based on the GHS [UN 2003], EPA [1998], and EU [2001] classification systems) of the test substances for each substance evaluated in this study. Using this information, combined with the results provided in the published literature, the accuracy and reliability of the tested

⁶ The prediction model noted to have the best performance contained three variables, TH10, TC10 (mean detection time for coagulation with a 10% solution), and IC₅₀. However, the prediction model shown in the reference only indicated two variables, TH10 (repeated twice) and IC₅₀. According to the text, it appears that one of the TH10 variables in the equation should be TC10; it is unclear from the text which TH10 should be changed to TC10.

⁷ The numbers used to generate these values were not provided in the literature study and no attempt was made to calculate the values from the prediction model provided in the reference.

version of the HET-CAM test method was determined and the results are provided in **Sections 6.0** and **7.0**.

In vivo and *in vitro* data was submitted by Dr. med. Horst Spielmann and Dr. Manfred Liebsch of ZEBET. Individual animal responses for the days that the animals were observed were provided. This data was used to identify the ocular irritant potential (based on the GHS [UN 2003], EPA [1998], and EU [2001] classification systems) of the test substances for each substance evaluated in this study. *In vitro* scores for each test substance for each individual testing laboratories was provided as were results using control substances. Using this information, the accuracy and reliability of the tested version of the HET-CAM test method was determined, and the results are provided in **Sections 6.0** and **7.0**. Results of the analyses conducted with control test substances are provided in **Section 2.0**.

In response to a request from NICEATM, *in vitro* HET-CAM scores were obtained for the substances evaluated in Balls et al. (1995). The data, provided by ECVAM, comprised of Q-Scores and S-Scores for all tested substances for each testing laboratory. The individual sample scores were not provided. Comparative *in vivo* individual rabbit data was obtained from the ECETOC database (ECETOC 1998). These data were used to identify the ocular irritant potential (based on the GHS [UN 2003], EPA [1998], and EU [2001] classification systems) of each substance evaluated. Using this information, combined with the results provided in the published literature, the accuracy and reliability of the tested version of the HET-CAM test method used in this study were determined, and the results are provided in **Sections 6.0** and **7.0**.

In vitro data was submitted by Dr. Philippe Vanparys and Dr. Freddy Van Goethem of Johnson & Johnson Pharmaceutical R&D (a division of Janssen Pharmaceutica N.V.). Times of development of endpoints for each egg tested for substances were provided for data presented in Gilleron et al. (1996, 1997). Furthermore, results from studies using control substances were provided upon request. Using this information, combined with the results provided in the published literature, the accuracy and reliability of this version of the HET-CAM test method used in this study were determined and the results of this re-analysis are provided in **Section 6.0** and **Section 7.0**. Results of the analyses conducted with control test substances are provided in **Section 2.0**.

[This Page Intentionally Left Blank]