



Published in final edited form as:

Regul Toxicol Pharmacol. 2011 October ; 61(1): 98–104. doi:10.1016/j.yrtph.2011.06.006.

Using Fewer Animals to Identify Chemical Eye Hazards: Revised Criteria Necessary to Maintain Equivalent Hazard Classification

Joseph K. Haseman^a, David G. Allen^b, Elizabeth A. Lipscomb^b, James F. Truax^b, and William S. Stokes^{c,*}

^aJ. K. Haseman Consulting, Raleigh, North Carolina 27614, USA

^bIntegrated Laboratory Systems, Inc., Research Triangle Park, North Carolina 27709, USA

^cThe National Institute of Environmental Health Sciences, The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, Research Triangle Park, North Carolina 27709, USA

Abstract

U.S. Federal Hazardous Substances Act (FHSA) regulations specify eye safety testing procedures and hazard classification criteria for substances regulated by the U.S. Consumer Product Safety Commission (CPSC). Current regulations require up to three sequential 6-animal tests. Testing consistent with the Organisation for Economic Cooperation and Development (OECD) test guideline for eye irritation/corrosion, which specifies 3 animals, can also be submitted to U.S. agencies. However, current FHSA regulations do not provide criteria to classify results from 3-animal tests. An analysis was conducted to determine criteria using results from 3-animal tests that would provide equivalent labeling to FHSA regulations. The frequency that FHSA requirements identify substances as ocular irritants was compared with the frequency that a criterion of either $\geq 1/3$ or $\geq 2/3$ positive animals would identify these substances. A database of rabbit eye tests was also used to estimate over- and underprediction rates for each criterion. In each instance, a criterion of $\geq 1/3$ positive animals more closely matched the expected outcome based on FHSA requirements, while a criterion of $\geq 2/3$ positive animals identified far fewer irritants. Using a classification criterion of $\geq 1/3$ positive animals provided equivalent or greater eye hazard labeling as current FHSA requirements, while using 50–83% fewer animals.

Keywords

CPSC; eye irritant; FHSA; hazard classification; hazard labeling; ocular safety testing

1. Introduction

Each year, approximately 2 million eye injuries occur in the U.S. (McGwin et al. 2006a). Of these, more than 40,000 result in permanent visual impairment. Household cleaning chemicals and other chemical products are the leading cause of consumer product-related

*Corresponding author: 530 Davis Drive, P.O. Box 12233, K2-16, Research Triangle Park, North Carolina 27709, USA, Phone: 1-919-541-7997, Fax: 1-919-541-0947, stokes@niehs.nih.gov.

Conflict of Interest Statement

The authors declare that there are no conflicts of interest.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

eye injuries in children under age 10 (McGwin et al. 2006b). In order to provide warnings to consumers and workers of the potential for chemicals and products to cause eye injuries, regulatory authorities require ocular safety testing to determine if substances may cause temporary or permanent eye damage. Testing results are then used for hazard classification and labeling of eye injury potential as required by appropriate national and/or international hazard classification systems. If classified as an eye hazard, hazard labeling of the chemical or product is required to warn users of the potential to cause temporary or permanent eye injuries, to provide the safety precautions necessary to avoid injuries, to provide the immediate first-aid procedures that should be followed in case of an accidental exposure, and to provide guidance on whether medical care should be sought.

The U.S. Federal Hazardous Substances Act (FHSA¹) “requires that certain hazardous household products (“hazardous substances”) bear cautionary labeling to alert consumers to the potential hazards that those products present and to inform them of the measures they need to protect themselves from those hazards” (see <http://www.cpsc.gov/businfo/fhsa.html>). The U.S. Consumer Product Safety Commission (CPSC) issues regulations implementing the FHSA. The regulations for hazardous substances under the FHSA are found in Title 16 Part 1500 of the U.S. Code of Federal Regulations (16 CFR 1500 [CPSC 2010]). Current U.S. ocular hazard classification regulations to implement FHSA labeling requirements for these products are provided in the Test for Eye Irritants (16 CFR 1500.42 [CPSC 2010]). This test provides criteria and procedures for identifying ocular hazards based on rabbit eye test results.

Current FHSA regulations require 6 animals per test and may require up to three sequential tests for each substance, thereby requiring 6, 12, or 18 animals to reach a hazard decision. The requirement for second and third sequential tests is based on the number of positive responses in the previous test. In 2002, the Organisation for Economic Co-operation and Development (OECD) adopted U.S. proposed revisions to Test Guideline 405: Acute Eye Irritation/Corrosion (OECD 2002) to reduce the maximum number of required animals from 6 to 3. Testing conducted in accordance with the OECD test guideline can be used to meet CPSC labeling requirements. However, current FHSA regulations do not provide criteria to classify results from a 3-animal test. Therefore, an analysis was conducted to determine classification criteria based on results from a 3-animal test that would provide hazard classification equivalent to that provided by current FHSA regulations, which require the use of 6 to 18 animals.

The National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) and the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) used the results from this analysis to develop recommendations for updating the CPSC Test for Eye Irritants (CPSC 2010) to require a maximum of 3 animals per test substance, which would be consistent with current ocular safety testing guidelines for the U.S. Environmental Protection Agency (EPA 1998) and the OECD (OECD 2002). ICCVAM recommendations are provided to Federal agencies to assist them in meeting Federal laws that require agencies, before adopting new alternative test methods, to determine that the test method will generate data in an amount and of a scientific value that is at least equivalent to the data generated from existing tests for hazard identification or risk assessment purposes (PHS 2000).

¹Abbreviations used: CFR, U.S. Code of Federal Regulations; CPSC, U.S. Consumer Product Safety Commission; EPA, U.S. Environmental Protection Agency; FHSA, U.S. Federal Hazardous Substances Act; ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; NICEATM, National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods; OECD, Organisation for Economic Co-operation and Development; PHS, Public Health Service

2. Testing requirements

The testing requirements necessary to determine the ocular hazard potential for substances regulated under the FHSA (FHSA 2008) are provided in 16 CFR 1500.42 (CPSC 2010) (see Table 1). Testing is conducted using an initial group of 6 albino rabbits, and 0.1 mL or 0.1 grams of the test substance is placed in the conjunctival sac of one eye with the contralateral eye serving as a negative or solvent control. Observations and severity scores are recorded at 24, 48, and 72 hours after test substance administration for four types of ocular injuries: corneal ulceration/opacity, conjunctival redness, conjunctival swelling, and iritis (see Table 2). Positive responses for individual animals are based on meeting or exceeding the minimum severity criteria for any one of the four types of eye injuries at any of the three time points. Criteria based on the number of positive animals are provided for each sequential test as to whether the hazard test result is positive, negative, or if a second or third test is required (Table 1).

The U.S. proposed revisions to OECD Test Guideline 405: Acute Eye Irritation/Corrosion (OECD 1987) to reduce the maximum number of required animals from 6 to 3 (deSilva et al. 1997; OECD 1999; Springer et al. 1993). The revised Test Guideline 405 was adopted in 2002 (OECD 2002). In accordance with the OECD Mutual Acceptance of Data Treaty (OECD 1981), U.S. agencies accept test data for review generated in accordance with OECD test guidelines.

The Animal Welfare Act (2010) requires that only the minimum number of animals necessary to obtain scientifically valid results be used, and the Public Health Service requires that a rationale for the appropriateness of the number of animals used be provided to and approved by the Institutional Animal Care and Use Committee (PHS 2002). In light of these policies and regulations, it is expected that most *in vivo* ocular safety testing would adhere to the 3-animal procedure described in the OECD and EPA test guidelines (OECD 2002; EPA 1998). However, current FHSA regulations do not provide criteria to classify results from a 3-animal test. Therefore, an analysis was conducted to determine classification criteria based on results from a 3-animal test that would provide hazard classification equivalent to that provided by current FHSA regulations that require the use of 6 to 18 animals.

3. Optimization of the number of positive animals required for FHSA hazard classification and labeling

The minimum number of animals that would be required under the FHSA sequential testing strategy to assign a definitive test classification as positive or negative was evaluated for each of the possible test outcomes (Table 3). The minimum percentage of positive animal responses that can result in a positive FHSA hazard classification is 22% ($2/6+1/6+1/6$ or $4/18$). The maximum percentage of positive animal responses that can result in a negative FHSA hazard classification is 17% ($1/6$) to 28% ($3/6+2/6+0/6$ or $5/18$) (Table 3). Ideally, a classification system should not produce internal inconsistencies, where the percentage of positive animal responses that can result in an irritant or not labeled hazard classification overlap.

4. Comparison of three strategies for reducing animal use for FHSA ocular hazard classification

The percentage of substances labeled as ocular irritants based on three different classification strategies was compared. The current sequential testing strategy used to assign an FHSA classification is denoted as Strategy 1. Strategy 2 represents a minimum threshold

of $\geq 1/3$ (33%) positive animals. Strategy 3 represents a minimum threshold of $\geq 2/3$ (67%) positive animals.

In order to compare the frequency with which each strategy would identify substances as ocular irritants, a number of different underlying population positive response rates were examined. This population positive response rate, denoted by p , is the overall likelihood that an animal will show a positive response for a given substance. Importantly, it is a “population” response rate, not the response rate observed in a given sample of 3 to 6 animals. However, for a specified value of p , it is possible to compute the likelihood of observing various responses in a given sample using binomial probabilities. This is illustrated in Table 4 for a general p , and for $p=20\%$ and $p=60\%$ to provide specific examples. For example, for a substance with an underlying positive response rate of $p=60\%$, the likelihood is 0.311 (31.1%) that there will be exactly 4 positive animals in a sample of 6 animals.

Table 5 presents the likelihood of classifying a substance as an ocular irritant for various underlying values of p . However, it does not show whether or not this classification is “correct” because this would require knowledge of the underlying positive response rate that differentiates irritants from nonirritants. As indicated in Table 3, the weakest possible response that is considered positive by the current sequential testing strategy is 22% (4/18), while a response of 17% (1/6 or 3/18) is considered negative. Therefore, it could be argued that the threshold positive response rate for considering a substance as an irritant for the current FHSA requirements should logically lie between 17% and 22%, perhaps 20%. However, this conclusion is complicated by the fact that an observed response rate of 28% (5/18) may occur and result in a chemical to not be labeled as an irritant (see Table 3). Because the underlying positive response rates in a population that are characteristic of an irritant or a nonirritant are not definitively known, a range of different underlying positive response rates were compared (Table 5) and presented graphically in Figure 1.

For purposes of illustration, consider $p=20\%$. Table 6 summarizes all the possible ways in which Strategy 1 could lead to a negative classification for a substance with a 20% population positive response rate. The probabilities in Table 6 are derived from Table 4. Thus, by subtraction from 1.0, the likelihood of a positive classification for Strategy 1 for $p=20\%$ is $1 - 0.796$ or 0.204 or 20.4% (see Table 5).

These calculations are much simpler for Strategies 2 and 3. The likelihood of a positive classification using Strategy 2, assuming $p=20\%$, is just the likelihood of observing 1/3, 2/3, or 3/3 positive responses, which using the probabilities in Table 4, are $0.384 + 0.096 + 0.008 = 0.488$ or 48.8% (see Table 5). For Strategy 3 and $p=20\%$, the likelihood of a positive classification is the sum of the likelihood of observing 2/3 or 3/3 positive responses, which is $0.096 + 0.008 = 0.104$ or 10.4% (see Table 5).

Even though it uses fewer animals, Strategy 2 is more powerful than current FHSA requirements for detecting positive response rates of up to 40% and has approximately the same power for response rates of 50% and greater (Figure 1). Strategy 3 will identify far fewer irritants than Strategy 2 for underlying positive response rates of 80% and fewer. Strategy 3 considers a single positive response (1/3) to not indicate an irritant response, and Strategy 3 has lower power than current FHSA requirements for underlying positive response rates of 20% to 80%.

The previous calculations were based on a variety of underlying positive response rates without consideration of whether or not they reflect the positive response rates seen in practice. Rather than assuming that each irritant and nonirritant has its own unique (and unknown) underlying positive response rate, a potentially useful approach is to derive a

mathematical model that accurately describes the observed distribution of positive responses seen for a large database of test substances. If a definitive structure can be imposed upon the data (and if the model fits the data), then the model parameters can be used to estimate over- and underprediction rates. With this in mind, a NICEATM database of 481 rabbit eye test studies using 6 animals each was analyzed. This database includes a wide range of chemical and product classes and represents the types of test substances typically evaluated in ocular safety testing (see Tables 7 and 8). Chemical classes were assigned to each substance using a standard classification scheme based on the National Library of Medicine Medical Subject Headings (MeSH®) classification system. If not assigned in the study report, the product class was sought from other sources, including the National Library of Medicine's ChemIDplus® database.

To calculate the estimated over- and underprediction rates for the three strategies using the NICEATM database, the first step was to find a model that fits the observed outcomes (see Table 9), some of which are irritants and some of which are nonirritants. We used a model that assumed a mixture of three binomial distributions, because it is unlikely that every irritant has exactly the same likelihood of producing a positive response in an animal. We assumed that the irritants could be categorized into two groups: Type I irritants (high underlying positive response rate) and Type II irritants (smaller underlying positive response rate).

From the observed distribution of positive animals in a 6-animal test, five key parameters were estimated: the underlying positive response rates for nonirritants and Type I and Type II irritants, and the percentage of Type I and Type II irritants in the database (the percentage of nonirritants in the database can then be calculated by subtraction from 100%). The following parameter estimates provided the best fit to our database:

Model parameter estimates for the NICEATM database:

- Type I irritants: Underlying positive response rate = 97.8%
- Type II irritants: Underlying positive response rate = 50.0%
- Nonirritants: Underlying positive response rate = 1.7%
- Percentage of Type I irritants in the sample: 54% or 260 substances
- Percentage of Type II irritants in the sample: 12.9% or 62 substances
- Percentage of nonirritants in the sample: 33.1% or 159 substances

Given this excellent fit to the data as indicated in Table 9, we calculated the percentage of substances that would be labeled as ocular irritants using each of the three strategies (see Table 10). The likelihood that a Type I irritant would be labeled as an ocular irritant is close to 100% for all three strategies. The likelihood that a Type II irritant would be labeled as an ocular irritant is approximately 88% for Strategies 1 and 2 but 50% for Strategy 3. The likelihood of labeling a nonirritant as an ocular irritant is 0% for Strategy 1, 5.0% for Strategy 2, and 0.1% for Strategy 3 (Table 10).

Based on these outcomes, the underlying over- and underprediction rates associated with this model were then calculated. All three strategies have a very low underprediction rate for Type I irritants. However, for Type II irritants, Strategies 1 and 2 have underprediction rates of approximately 12%, while Strategy 3 has a 50% underprediction rate. For nonirritants, Strategies 1 and 3 have very low overprediction rates, while the overprediction rate for Strategy 2 is 5% (see Table 11).

It is important to note that this approach is similar to the approach used by Springer et al. (1993) except for the fact that we assumed two different underlying positive response rates for irritants, whereas Springer et al. used only one (i.e., they assumed that every irritant has exactly the same likelihood of producing a positive response in an animal). Based on the distribution of positive animals in a 6-animal test in the NICEATM database, the use of two different underlying positive response rates for irritants provided a much better fit to the data.

5. Previous proposals to reduce the number of animals used for ocular safety testing

Results from DeSousa et al. (1984) and Talsma et al. (1988) showed that using 3 rabbits per test provided accuracy of up to 94% in predicting a 6-animal test (using subsets of 3 animals). Springer et al. (1993) also conducted analyses to determine if the standard group size of 6 rabbits for ocular safety testing could be reduced in order to use fewer animals and concluded that a 3-animal test and a decision rule requiring at least 2 positive animals to classify a substance as an irritant yielded accuracy of 98%.

As indicated above, the model used by Springer et al. (1993) assumed two mutually exclusive populations: irritants and nonirritants, each population having a single underlying positive response rate estimated from the data. They fit a mixture of two binomial models to each of four different databases, but the only database with a distribution of outcomes that closely matched the NICEATM database of 481 rabbit eye test studies was an EPA database of 48 substances. Springer et al. (1993) reported the following parameter estimates for the EPA database:

- Irritants: Underlying positive response rate = 95.0%
- Nonirritants: Underlying positive response rate = 8.6%
- Percentage of nonirritants in the sample: 35%
- Percentage of irritants in the sample: 65%

Note that the estimated percentage of nonirritants in the EPA database (35%) is very similar to our own estimate (33.1%) for the much larger NICEATM database, but the Springer et al. model does not differentiate between Type I and Type II irritants. As a result, their parameter estimates provided a poor fit to the NICEATM database of 481 studies (Table 12). In fact, we found that their model did not provide a good fit to the EPA data upon which their parameter estimates were based (e.g., predicting only 0.2 3/6 outcomes compared with 3 actually observed, a 15-fold underprediction). This lack of model fit was more apparent using the NICEATM database of 481 substances, which was approximately 10-fold larger than the Springer et al. (1993) EPA database.

The largest database used by Springer et al. (1993) was the 139-substance Marzulli and Ruggles database, but the pattern of response seen in these studies was quite different from that seen in the NICEATM database of 481 studies. Even so, the best-fitting Springer et al. (1993) model showed the same lack of fit problem. For example, ten 3/6 positive responses were observed compared with only 3.1 predicted by the best-fitting Springer et al. (1993) model.

It is important to understand the factors that led to different conclusions in our evaluation, which favored Strategy 2, and that of Springer et al. (1993), which favored Strategy 3. For example, Table 1 in Springer et al. (1993) suggests that Strategy 2 may have an unacceptably high overprediction rate.

The primary reason for the different conclusions is that the EPA 48-substance database was of insufficient size to detect the Type II irritants that were producing positive response rates of approximately 50%. By not taking these irritants into account, the Springer et al. (1993) model underestimated the underprediction rate for Strategy 3, because this strategy does not perform well for detecting positive response rates of approximately 50% (see Table 5).

Another consequence of Springer et al. (1993) ignoring the Type II irritants was a 5-fold overestimation of the positive response rate of nonirritants (8.6% vs. 1.7%). This difference is important because the overprediction rate of Strategy 2 increases substantially as the assumed positive response rate for nonirritants increases (see Table 5). It is the Springer et al. (1993) overestimation of the positive response rate for nonirritants that produced the artificially high overprediction rate for Strategy 2 shown in their Table 1.

6. Conclusion

The results indicate that using a classification criterion of at least one out of three positive animals in a 3-animal test for the identification of eye hazards will provide the same or greater level of eye hazard labeling as current FHSA requirements, while using 50% to 83% fewer animals. A criterion of at least two out of three positive animals in a 3-animal test will identify far fewer irritants, especially those irritants with a smaller underlying positive response rate. Accordingly, this analysis should facilitate regulatory decisions on classification criteria that will support the adoption of test methods using fewer animals. The analysis is also expected to assist agencies in complying with U.S. laws requiring that, before adopting alternative methods, that they determine that the test method will generate data in an amount and of a scientific value that is at least equivalent to the data generated from existing tests for hazard identification or risk assessment purposes.

Acknowledgments

Funding

This work was supported by the Intramural Research Program of the National Institutes of Health, National Institute of Environmental Health Sciences. ILS staff are supported by NIEHS contract N01-ES 35504. The views expressed in this manuscript do not necessarily represent the official position of any U.S. Federal agency.

References

- Animal Welfare Act, 2010 (amended). Public Law 89–544, 7 U.S.C. 2131 et seq.
- ChemIDplus. Bethesda, MD: National Library of Medicine; Available: <http://chem.sis.nlm.nih.gov/chemidplus/>
- CPSC. Federal Hazardous Substances Act Regulations, 16 CFR 1500. 2010.
- deSilva O, et al. Evaluation of eye irritation potential: statistical analysis and tier testing strategies. *Food Chem Toxicol.* 1997; 35:159–164. [PubMed: 9100818]
- DeSousa D, et al. Statistical consequences of reducing the number of rabbits utilized in eye irritation testing: Data on 67 petrochemicals. *Toxicol Appl Pharmacol.* 1984; 76:234–242. [PubMed: 6495331]
- EPA. Health Effects Test Guideline, OPPTS 870.2400 Acute Eye Irritation. U.S. Environmental Protection Agency; Washington, DC: 1998.
- FHSA, 2008 (amended). Public Law 86–613. 15 U.S.C. 1261–1278.
- McGwin G, et al. *Invest Ophthalmol Vis Sci.* 2006a; 47:521–527. [PubMed: 16431945]
- McGwin G, et al. *J Safety Res.* 2006b; 37:501–506. [PubMed: 17113601]
- MeSH (Medical Subject Headings). National Library of Medicine; Bethesda, MD: Available: <http://www.nlm.nih.gov/mesh>

- OECD. Council Act, Mutual Acceptance of Data. Organisation for Economic Cooperation and Development; Paris: 1981.
- OECD. Test Guideline 405, Acute Eye Irritation/Corrosion. Organisation for Economic Co-operation and Development; Paris: 1987.
- OECD. OECD Series on Testing and Assessment, Number 14. ENV/JM/MONO(99)4. Organisation for Economic Co-operation and Development; Paris: 1999. Detailed Review Document on Classification Systems for Eye/Irritation/Corrosion in OECD Member Countries.
- OECD. Test Guideline 405, Acute Eye Irritation/Corrosion. Organisation for Economic Co-operation and Development; Paris: 2002.
- PHS, 2000. 42 U.S\, 289d. Subchapter III–National research institutes; Part H–General provisions; Section 289d–Animals in research.
- PHS. Public Health Service Policy on Humane Care and Use of Laboratory Animals. Office of Laboratory Animal Welfare, National Institutes of Health, Department of Health and Human Services; Bethesda, MD: 2002.
- Springer J, et al. Number of animals for sequential testing. *Food Chem Toxicol.* 1993; 31:105–109. [PubMed: 8449454]
- Talsma D, et al. Reducing the number of rabbits in the Draize eye irritancy test: A statistical analysis of 155 studies conducted over 6 years. *Fundam Appl Toxicol.* 1988; 10:146–153. [PubMed: 3350224]

Highlights

- FHSA regulations currently specify a classification system based on sequential 6-animal tests
- Current best practices for eye irritation/corrosion testing involve testing of up to a total of 3 animals.
- We developed criteria for FHSA hazard classification based on results from a 3-animal test
- 481 rabbit eye test studies were used to calculate the estimated over- and underprediction rates
- Our results indicate that $\geq 1/3$ positive animals approximates the level of eye hazard labeling from current FHSA requirements

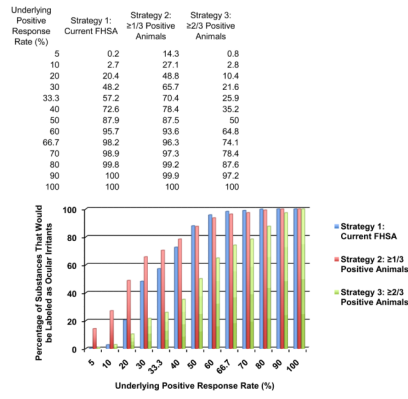


Figure 1.
 Strategy 2 provides the same or greater level of eye hazard labeling as current FHSA requirements
 Intended for color reproduction on the Web and in print.

Table 1

FHSA classification criteria and procedures for identification of ocular hazards (16 CFR 1500.42)

<p>Criteria for a Positive Response in a Single Rabbit (Based on one or more of the following for any reading at 24, 48, and 72 hours)</p>	<ol style="list-style-type: none"> 1 Corneal ulceration (other than a fine stippling) or corneal opacity^a ≥1 2 Iritis^b ≥1 3 Conjunctival swelling^c ≥2 4 Conjunctival redness^c ≥2
<p>Hazard Classification Criteria and Procedures</p>	<p><u>First Test: Test 6 animals</u></p> <ul style="list-style-type: none"> • If ≥4/6 animals are positive, the test is positive. • If ≤1 animal is positive, the test is negative. • If 2/6 or 3/6 animals are positive, a second test is conducted using a different group of 6 animals. <p><u>Second Test: Test 6 animals</u></p> <ul style="list-style-type: none"> • If ≥3/6 animals are positive, the test is positive. • If 0/6 are positive, the test is negative. • If 1/6 or 2/6 is positive, a third test is conducted using a different group of 6 animals. <p><u>Third Test: Test 6 animals</u></p> <ul style="list-style-type: none"> • If ≥1/6 animals are positive, the test is positive. • If 0/6 are positive, the test is negative.

FHSA = Federal Hazardous Substances Act (Public Law 86–613, 16 CFR 1500)

^aUlceration of the cornea (other than a fine stippling) or opacity of the cornea (other than a slight dulling of the normal luster)

^bInflammation of the iris (other than a slight deepening of the folds [or rugae] or a slight circumcorneal injection of the blood vessels)

^cObvious conjunctival swelling with partial eversion of the lids or conjunctival redness with diffuse crimson-red; individual vessels not easily discernible

Table 2

Scores for grading severity of ocular lesions

Lesion ^a	Score
Cornea	
Scattered or diffuse areas of opacity (other than slight dulling of normal luster), details of iris clearly visible	1
Easily discernible translucent areas, details of iris slightly obscured	2
Opalescent areas, no details of iris visible, size of pupil barely discernible	3
Complete corneal opacity, iris not discernible	4
Iris	
Markedly deepened folds, congestion, swelling, moderate circumcorneal injection (any one of these or combination of any thereof), iris still reacting to light (sluggish reaction is positive)	1
No reaction to light, hemorrhage, gross destruction (any one or all of these)	2
Conjunctiva	
A. Redness (refers to palpebral and bulbar conjunctiva only)	
Some vessels definitely injected above normal	1
Diffuse, crimson red, individual vessels not easily discernible	2
Diffuse beefy red	3
B. Chemosis	
Any swelling above normal (includes nictitating membrane)	1
Obvious swelling with partial eversion of the lids	2
Swelling with lids about half closed	3
Swelling with lids about half closed to completely closed	4

Table is adapted from the *CPSC Illustrated Guide for Grading Eye Irritation Caused by Hazardous Substances*, available by written request: Directorate for Epidemiology and Health Sciences, CPSC, Washington, DC 20207.

Scores in **bold** indicate positive responses. Scores of 0 are assigned for each parameter if the cornea, iris, or conjunctiva is normal.

^aPositive responses for individual animals are based on meeting or exceeding the minimum severity criteria for any one of the four types of eye injuries at any of the three time points.

Table 3

Number of positive animals and sequential tests required for assignment of an irritant classification according to the current FHSA requirements

Positive Test Criteria for Irritant Classification:	Positive Animals					
	≥4/6	2/6 or 3/6	3/6	3/6	2/6	2/6
First Test Results			3/6	3/6	2/6	2/6
Results from Second Test (when required)	Second test not required	≥3/6	2/6	1/6	2/6	1/6
Results from Third Test (when required)	Third test not required	Third test not required	≥1/6	≥1/6	≥1/6	≥1/6
Minimum Number of Positive Animals for Irritant Classification	4/6 (67%)	5/12 (42%)	6/18 (33%)	5/18 (28%)	5/18 (28%)	4/18 (22%)
Maximum Number of Positive Animals for Not Labeled Classification	1/6 (17%)	3/12 (25%)	5/18 (28%)	4/18 (22%)	4/18 (22%)	3/18 (17%)

FHSA = Federal Hazardous Substances Act (Public Law 86-613, 16 CFR 1500)

Probability of observing 0 to 6 positive animals in a sample of n=3 or n=6 for various population positive response rates (p) assuming a binomial model

Table 4

No. Positive Animals in a Sample	Probability of Response in Sample		Probability of Response in Sample			
	n=3	n=6	n=3 p=20%	n=3 p=60%	n=6 p=20%	n=6 p=60%
0	$(1-p)^3$	$(1-p)^6$	0.512	0.064	0.262	0.004
1	$3p(1-p)^2$	$6p(1-p)^5$	0.384	0.288	0.393	0.037
2	$3p^2(1-p)$	$15p^2(1-p)^4$	0.096	0.432	0.246	0.138
3	p^3	$20p^3(1-p)^3$	0.008	0.216	0.082	0.276
4	-	$15p^4(1-p)^2$	-	-	0.015	0.311
5	-	$6p^5(1-p)$	-	-	0.002	0.187
6	-	p^6	-	-	<0.001	0.047

Table 5

Percentage of substances labeled as ocular irritants based on various population positive response rates (p) for the three strategies

Population Positive Response Rate (p)	Percentage of Substances That Would be Labeled as Ocular Irritants		
	Strategy 1 Current FHSA	Strategy 2 $\geq 1/3$ Positive Animals	Strategy 3 $\geq 2/3$ Positive Animals
1.7% ^a	0.0%	5.0%	0.1%
5%	0.2%	14.3%	0.8%
10%	2.7%	27.1%	2.8%
20%	20.4%	48.8%	10.4%
30%	48.2%	65.7%	21.6%
33.3%	57.2%	70.4%	25.9%
40%	72.6%	78.4%	35.2%
50% ^a	87.9%	87.5%	50.0%
60%	95.7%	93.6%	64.8%
66.7%	98.2%	96.3%	74.1%
70%	98.9%	97.3%	78.4%
80%	99.8%	99.2%	87.6%
90%	100%	99.9%	97.2%
97.8% ^a	100%	100%	99.9%
100%	100%	100%	100%

^aEstimated underlying positive response rates for the NICEATM database (see Table 9)

Table 6

Probability that Strategy 1 will result in a negative classification for p=20%

Strategy 1 Test Result			Probability
Test 1	Test 2	Test 3	
0/6	-	-	0.262
1/6	-	-	0.393
2/6	0/6	-	$(.246)(.262) = .0645$
3/6	0/6	-	$(.082)(.262) = .0215$
2/6	1/6	0/6	$(.246)(.393)(.262) = .0253$
3/6	1/6	0/6	$(.082)(.393)(.262) = .0084$
2/6	2/6	0/6	$(.246)(.246)(.262) = .0159$
3/6	2/6	0/6	$(.082)(.246)(.262) = .0053$
Total	-	-	0.796

Table 7

Chemical classes in the NICEATM database

Chemical Class	No. of Substances	Chemical Class	No. of Substances
Alcohols	40	Isocyanates	4
Aldehydes	2	Lactones	1
Aluminum compounds	1	Nitriles	2
Amidines	1	Nitro compounds	4
Amines, including salts	26	Oniums	4
Anhydrides	1	Organics, including salts	19
Boron compounds	1	Organometallic compounds	1
Carboxylic acids, including salts	9	Organophosphorus compounds	6
Esters	26	Organosilicon compounds	10
Ethers	43	Peroxides	1
Formulations	215	Phenols	9
Heterocycles	10	Polycyclic compounds	3
Hydrocarbons (acyclic, cyclic, halogenated)	35	Quinones	2
Imides	1	Sulfur compounds	9
Inorganic, including salts	8	Urea	1

Table 8

Product classes in the NICEATM database

Product Class	No. of Substances	Product Class	No. of Substances
Adhesives	2	Inks	4
Algicides	1	Insect repellents	8
Antifreezes	2	Liquid nylon products	1
Antiperspirants, roll-on	1	Lubricants	4
Bleaches	1	Oils	2
Bonding agents	2	Pesticides	22
Caustic agents	1	Petroleum products	9
Chemical intermediates	1	Pharmaceutical products, metabolites, or intermediates	3
Cleaners	11	Plasticizers	5
Cosmetics (creams, lotions, wetting agents, etc.)	41	Plastics additives	15
Detergents	20	Resins	3
Dyes	3	Rodenticides	1
Fabric softeners	1	Shampoos and conditioners, hair	14
Flame retardants	15	Solvents	17
Flocculating agents	1	Sulfurs	1
Floor strippers	1	Sunscreens	4
Fungicides	5	Surfactants (anionic, cationic, nonionic)	44
Germicides	6	Textiles	1
Heat transfer fluids	1	Thermal paper coatings	1
Industrial chemicals	6	-	-

Table 9
 Goodness of fit for a database of 481 test results using a mixture of three binomial distributions

Number of Positive Animals in a 6-Animal Test	Predicted Type I Irritants	Predicted Type II Irritants	Predicted Nonirritants	Total Predicted by NICEATM Model	Observed in NICEATM Database
0	0	1.0	143.4	144.4	142
1	0	5.8	15.0	20.8	21
2	0	14.5	0.6	15.1	19
3	0.1	19.4	0	19.5	15
4	1.7	14.5	0	16.2	20
5	30.7	5.8	0	36.5	35
6	227.5	1.0	0	228.5	229
Total	260 (54.0%)	62 (12.9%)	159 (33.1%)	481	481

Table 10

Percentage of substances labeled as ocular irritants based on estimated underlying positive response rates for three strategies: three binomial distributions

Three Binomial Distributions	Estimated Underlying Positive Response Rate	Percentage of Substances That Would be Labeled as Ocular Irritants		
		Strategy 1 Current FHSA	Strategy 2 $\geq 1/3$ Positive Animals	Strategy 3 $\geq 2/3$ Positive Animals
Nonirritants	1.7%	0%	5.0%	0.1%
Type II Irritants	50%	87.9%	87.5%	50.0%
Type I Irritants	97.8%	100%	100%	99.9%

Table 11

Percentage of substances that would be over- and underpredicted for the three strategies

Three Binomial Distributions	Strategy 1 Current FHSA	Strategy 2 $\geq 1/3$ Positive Animals	Strategy 3 $\geq 2/3$ Positive Animals
Percentage of Substances That Would be Overpredicted			
Nonirritant	0%	5.0%	0.1%
Percentage of Substances That Would be Underpredicted			
Type II Irritants	12.1%	12.5%	50.0%
Type I Irritants	0%	0%	0.1%

Table 12

Lack of fit using the Springer et al. (1993) model

Number of Positive Animals in a 6-animal Test	Predicted Irritants	Predicted Nonirritants	Total Predicted by Springer Model	Observed in NICEATM Database
0	0	98.2	98.2	142
1	0	55.5	55.5	21
2	0	13.0	13.0	19
3	0.7	1.6	2.3	15
4	9.5	0.1	9.6	20
5	72.6	0	72.6	35
6	229.8	0	229.8	229
Total	312.6 (65%)	168.4 (35%)	481	481