Same-Chemical Comparison of Nonanimal Eye Irritation Test Methods: Bovine Corneal Opacity and Permeability, EpiOcular™, Isolated Chicken Eye, Ocular Irritection®, OptiSafe™, and Short Time Exposure

Stewart Lebrun[1,*], **Linda Nguyen**[1], **Sara Chavez**[1], **Roxanne Chan**[1], **Debby Le**[1], **Minh Nguyen**[1], **James V. Jester**[2]

[1]Lebrun Labs LLC, Anaheim, California

[2]Department of Ophthalmology and Biomedical Engineering, University of California Irvine, Irvine, California

Abstract

The testing and classification of chemicals to determine adverse ocular effects are routinely conducted to ensure that materials are appropriately classified, labeled, and meet regulatory and safety guidelines. We have performed a same-chemical analysis using publicly available validation study results and compared the performance between tests for the same chemicals. To normalize for chemical selection, we matched chemicals tested by pairs of tests so that each matched set compared performance for the exact same chemicals. Same-chemical accuracy comparisons demonstrate a chemical selection effect that results in a wide range of overlapping false-positive (FP) rates and accuracies for all test methods. In addition, the analysis suggests that a tiered-testing strategy with specific combinations of tests can reduce the FP rate for some combinations. However, reductions in the FP rates were typically accompanied by an increase in the false-negative rates, resulting in minimal advantage in terms of accuracy. In addition, actual improvements in the FP rate after retesting positives with a second test are not as good as the theoretical improvements because some chemicals and functional groups appear to be broadly misclassified by all test methods, which, to the extent the tests make the same-chemical misclassifications, reduces the advantage of using tiered-testing strategies.

Keywords

ocular irritation; in vitro eye irritation tests; functional group; GHS

---

*Corresponding author at: Lebrun Labs LLC, 3301 E. Miraloma Ave., Ste. 194, Anaheim, CA 92806, SJ@lebrunlabs.com.

Introduction

Ocular irritancy testing plays an important role in the labeling of chemicals and products to protect consumers and manufacturers from exposure to toxic chemicals through appropriate product and chemical labeling. In the past, the appropriate labeling of chemicals and compounds has relied on the in vivo Draize rabbit eye test, which has become the "gold standard" for chemical classification. This live animal test uses a clinical scoring system that grades the severity and duration of the ocular irritation response that generally occurs from 1 to 21 days after a chemicals exposure. The resulting clinical scores and durations of effects are then applied to either the United Nations Globally Harmonized System (GHS; Choksi et al., 2020; Lebrun et al., 2019) or U.S. Environmental Protection Agency (EPA) methods of eye irritation classification (Choksi et al., 2020). Because of animal cruelty concerns, there has been a major effort over the past several decades to develop alternative strategies (non-live animal tests) that could be safely used to classify and label potentially harmful chemicals. In this paper, we briefly review the background concerning ocular irritation testing and chemical/compound labeling, the major driving forces for nonanimal testing, and the currently recognized and validated nonanimal tests; we then perform a same-chemical comparison analysis using published databases to compare accuracy for nonanimal eye irritation tests alone or in a two-test tier.

Classification of Ocular Irritation

The GHS classification system is based on the most severe effects in two-thirds of animals tested (typically three) when the score drives the classification. However, for severe, persistent injury, one animal drives the classification. Class definitions include not classified (NC, no serious eye damage averaged over the first three days), 2 (2A, reversible serious irritation effects that reverse by 21 days, or 2B, by 7 days), and 1 (extreme and/or irreversible irritation effects that do not reverse by 21 days) (Choksi et al., 2020). GHS classification may be mechanistically related to materials that cause a depth of eye injury that damages the basement membrane and corneal stroma, and therefore is a more serious injury than damage to the superficial epithelium and takes several days or longer to reverse (Lebrun et al., 2019). This serious level and duration of eye injury required for GHS classification is used by regulatory agencies to indicate when labeling for eye protection is required. Hence, the GHS classification "NC" does not indicate a material is a "nonirritant" as defined by the layperson including users of eye area cosmetics and personal care products.

The EPA classification system is based on the most severe animal responses (typically one out of three animals) with class definitions of IV (no significant damage 24 hours after exposure), III (damage reversible by 7 days after exposure), II (damage reversible by 21 days after exposure), and I (corrosive, lesions do not reverse by 21 days) (Choksi et al., 2020). While EPA category IV indicates no serious damage at 24 hours or longer, significant adverse ocular effects can still occur for these materials as long as these effects resolve prior to 24 hours. Hence, for some testing applications, for example, consumer satisfaction of eye

area products, additional classifications schemes with a shorter-duration sensitivity may be desirable.

The International Transition to In Vitro Eye Tests for Routine Product Testing

Over public concern regarding animal cruelty and other issues of live animal ocular irritation tests, the EU was the first to mandate nonanimal testing for routine product testing in 2013 under Cosmetics Regulation 1223/2009, including eye area cosmetics (European Community, 2009). In 2018, the EU Parliament urged a worldwide ban on testing cosmetics on animals by 2023. Currently, there are bans or restrictions on the use of animals for product testing or the sale of products tested on animals in about 40 countries worldwide (Grum, 2019). The percent of products tested using nonanimal testing methods is increasing, according to the most recent ECHA report under REACH regulation 1907/2006.

While the United States has been slow to legislate against live animal testing, statewide bans on the sale or import of cosmetics tested on animals started 2020 in California, Illinois, and Nevada (ALTEX, 2019). A bipartisan bill introduced in the U.S. 116th Congress substantially restricts the use of animals for testing for cosmetics, and sale and transport of animal-tested cosmetics with limited exceptions and stiffer penalties ("Humane Cosmetics Act of 2019," S.2886/H.R.5141, 2019). The bill requires the U.S. Food and Drug Administration to develop and publish on its website, within one year of the bill's enactment, a strategic plan "to the maximum extent possible with available resources, prioritize and carry out performance assessment, validation, and translational studies to accelerate the development of scientifically valid test methods and strategies that replace the use of vertebrate animals," and provide an updated list of alternative testing methods ("Humane Cosmetics Act of 2019," S.2886/H.R.5141, 2019).

Alternative Tests to Live Animal Testing and Their Validation

In vitro test methods have been developed and validated for the detection of ocular corrosives and GHS NC ocular irritants. In accordance with the move to not perform new live animal tests, retrospective animal test databases are commonly used as sources of in vivo data to determine the accuracy of nonanimal test methods in these validation studies. These databases include the Draize Eye Test Reference Database (DRD), ECHA chemical dossiers/REACH legislation (Luechtefeld et al., 2016b), Cosmetics Europe [a compilation of data sources, including include ECETOC, ZEBET, LNS, NICEATM, and European Centre for the Validation of Alternative Methods (ECVAM)] (Barroso et al., 2017), and ECETOC (1992; 1998).

However, it should be noted that there is variability of the animal data that places limits on the predictive performance of nonanimal test methods and makes validation of nonanimal tests more difficult. Specifically, in vivo databases often include single three-animal studies that are prone to variability, as demonstrated by the typical lack of agreement between responses within a given study. There are also reproducibility issues, as demonstrated by retrospective pooled analyses (Weil and Scala, 1971). In addition to the variability issues

associated with small group size, the inability of the clinical scoring systems to reflect the complexities of the total in vivo response are a major limitation (York and Steiling, 1998). There are also factors related to the dosing of test materials, methods of exposure, and the subjectivity of observations, scoring, and laboratory procedures (OECD, 1998). The Draize eye test has been found to demonstrate high misclassification errors. About 12% of the chemicals classified as category 2 and at least 11% of those classified as category 1 could in fact be equally identified as NC and category 2 (respectively) by the in vivo Draize eye test; based only on within-test variability (Luechtefeld et al., 2016a; Barroso et al., 2017).

Key criteria for selecting reference chemicals include chemicals covering different drivers of classification based on observed tissue effects (primarily corneal opacity), relevant chemical classes, and physical states. According to Barroso et al. (2017), "Considering the chemicals in the DRD that are commercially available today (511 individual chemicals tested in 556 studies), only about 73% (375 individual chemicals tested in 402 studies) are considered good reference chemicals that can be selected for future studies." Most recently, a set of 80 reference chemicals was selected in collaboration with Cosmetics Europe from their database of 634 chemicals that was generated from past validation studies and correlated with the main chemical ocular effect driving the in vivo GHS/EU CLP classification; these 80 chemicals consist of a wide range of classes used for in vitro eye test development for the CEFIC-LRI-AIMT6-VITO CON4EI (CONsortium for in vitro Eye Irritation testing strategy) project (Adriaens et al., 2018a).

In Vitro Test Methods

Validated and widely recognized tests used to detect GHS category 1 chemicals (ocular corrosives) include Bovine Corneal Opacity and Permeability (BCOP), Isolated Chicken Eye (ICE), Ocular Irritection, OptiSafe, and Short Time Exposure (STE); however, due to the complexity of the analysis and space constraints, GHS category 1 is not analyzed further in the current review. Validated nonanimal tests to detect GHS NC analyzed here include BCOP, EpiOcular, ICE, Ocular Irritection, OptiSafe, and STE, whose methods are briefly reviewed below. Data sources for each test method are listed in Table 1.

Bovine Corneal Opacity and Permeability

The BCOP method uses cow eyes from the meat industry to measure corneal opacity with an opacitometer and permeability using a spectrophotometer or plate reader. The BCOP test has been validated by the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) along with ECVAM and the Japanese Center for the Validation of Alternative Methods (JaCVAM) in 2006 and 2010, respectively, using the BCOP (OP-KIT) to identify materials that cause eye corrosion (GHS category 1) and materials not requiring classification for serious eye damage (GHS NC). The OECD test guideline was updated in 2020 to include use of a laser light-based opacitometer (LLBO), which demonstrated similar performance as the OP-KIT opacitometer (OECD 437, 2020a). The OP-KIT opacitometer is a center-weighted reading of white polychromatic light transmission, and the LLBO opacity reader measures the entire corneal surface with a monochromatic laser source. Permeability is quantified using optical density (OD) values of the amount of sodium fluorescein dye

that penetrates all layers. Using the opacity measurements and permeability OD values, an in vitro irritancy score (IVIS) is generated. A prediction model differentiates these scores into GHS NC and category 1 predictions. The OECD BCOP guideline states for the BCOP (OP-KIT) method that the accuracy was 69%, the false-positive (FP) rate was 69%, and the false-negative (FN) rate was 0%. For the BCOP (LLBO) method, the accuracy was 83%, the FP rate was 45%, and the FN rate was 6% (OECD 437, 2020a). The BCOP method can test solids and liquids, but is not recommended for use with test chemicals that are classified as irritating to eyes (category 2B or 2A) due to a number of category 1 chemicals underclassified and a number of NC chemicals overclassified (OECD 437, 2020a).

Reconstituted Human Corneal Epithelium

There are four validated (OECD 492, 2019a), commercially available, reconstituted human cornea-like epethelium (RhCE) test methods: EpiOcular™ Eye Irritation Test (EIT) (MatTek Corporation, Ashland, MA, USA) (Kaluzhny et al., 2011; Pfannenbecker et al., 2013), SkinEthic™ Human Corneal Epithelium (HCE) (L'Oreal, EpiSkin, France) EIT (Alepee et al., 2013; EURL ECVAM, 2014), Labcyte CORNEA-MODEL24 EIT (Katoh et al., 2013), and MCTT HCE™ EIT (Yang et al., 2017; Lim et al., 2019). These test methods are three-dimensional models that consist of epithelium composed of human epidermal keratinocytes (skin cells) or transformed corneal epithelial cells. The test matrix is constructed to model the corneal epithelium with stratified and noncornified cells (Doucet et al., 2006; Jung et al., 2011). The RhCE data used for this analysis include only those for EpiOcular EIT.

Ocular irritancy is predicted based on in vitro cell viability following exposure to the test chemical as determined by the MTT viability assay. Live cell metabolism turns MTT purple. The optical density of extracted purple MTT is quantified using a plate reader. A prediction model classifies substances as nonirritants based on the percent of treatment group cell viability compared to that of the control group. EpiOcular EIT distinguishes test substances not requiring classification for serious eye damage/eye irritancy (NC) from chemicals requiring classification and labeling, according to the GHS. The EpiOcular OECD guideline states that the accuracy was 80%, the FP rate was 37%, and the FN rate was 4% (OCED 492, 2019a). While a wide range of substances can be tested using EpiOcular EIT, colored substances and substances that directly change the color of MTT require additional controls (OCED 492, 2019a).

Isolated Chicken Eye

The ICE test method (OECD 438, 2018) is an organotypic model that uses chicken eyes from the meat industry to classify GHS NC and category 1 chemicals. Effects on the cornea are measured using a visual clinical scoring of opacity and involve applying fluorescein to the cornea, scoring damage to the epithelium, and measuring increased thickness (swelling) and macroscopic damage to the surface of the eye. Corneal opacity, swelling, and damage are assessed individually, and then all three endpoints are combined and applied to a prediction model to make an irritancy classification. The OECD ICE test guideline states that the accuracy was 88%, the FP rate was 24%, and the FN rate was 3% (OCED 438, 2018). While the ICE method can be used to test a wide range of substances, testing of

alcohols produced a high FP rate, and solids/surfactants produced a high FN rate (OCED 438, 2018).

Ocular Irritection

Ocular Irritection® (InVitro International, Irvine, CA, USA) is a macromolecular test used to identify GHS NC and category 1 chemicals. The test system is a cell-free matrix that mimics the cornea. This is used to measure the amount of coagulation and saponification following exposure to a test chemical (Eskes et al., 2014). Quantification of changes to the test matrix are converted into a maximal qualified score and a GHS classification result. The OECD Ocular Irritection test guideline indicates that the accuracy was 75%, the FP rate was 41%, and the FN rate was 9% (OCED 496, 2019b). Ocular Irritection can test solids and liquids when a 10% solution of those materials falls within the pH range between 4 and 9 (OECD 496, 2019b). The OECD application domain is limited to weight of evidence testing versus as a standalone test for GHS NC. Therefore, FN and FP rates "in this context are not critical since all test chemicals that come out negative or positive would be subsequently tested with other adequately validated in vitro test(s), using a sequential testing strategy in a weight-of-evidence approach according to the OECD GD 263" (OECD 496, 2019b).

OptiSafe

The OptiSafe™ (Lebrun Labs LLC, Anaheim, CA, USA) ("Optimized for Safety") test is a shelf-stable, test tube-based in chemico method. OptiSafe is used to determine whether an unknown substance is an ocular nonirritant or ocular corrosive using chemical models of three variables: 1) damage to the corneal stroma, 2) damage to phospholipid bilayers, and 3) the potential to induce pH extremes in a system (pH buffering system of the eye) (Choksi et al., 2020). OD measurements and pH are used to calculate a score that is applied to prediction models for GHS NC and GHS category 1 (ocular corrosive). It also includes EPA category IV and EPA category I (ocular corrosive) prediction models. A validation study coordinated by NICEATM with members of ICCVAM forming the validation management team, assessed accuracy for the GHS NC and EPA IV nonsurfactant chemical predictions. For the three-lab masked transferability phase, when results across all three laboratories were combined based on the majority classification for the GHS system, test method accuracy was 89%, the FN rate was 0%, and the FP rate was 23%. Additional chemicals were selected by the validation management team to evaluate the application domain; these were tested by the lead lab only in the "application domain phase." Based on results from both the masked transferability study and the masked application domain study conducted by the lead lab only, test method accuracy for the GHS system was 80% (when the degraded chemical iso-octylthioglycolate, CASRN 25103–09–7 is accounted for, see Supplemental Material I), the FN rate was 0%, and the FP rate was 40% (GHS). The application domain includes soluble and insoluble solids, liquids, surfactants, and highly colored substances.

Short Time Exposure (STE)

The Statens Seruminstitut Rabbit Cornea cell-based STE test method (OECD 491, 2020b) is an in vitro cytotoxicity test. Cell viability is measured after a 5-minute exposure at both 5% and 0.05% test material concentrations. While the cell line is not differentiated, this is balanced by the rapid exposure of cells to the test substance followed by an evaluation of

viability using the MTT assay (ICCVAM-NICEATM, 2013). The prediction model uses cell viability at both the 5% and 0.05% concentrations to classify GHS NC and GHS category 1 chemicals. No predictions can be made for GHS category 2B and 2A chemicals. For the detection of GHS NC chemicals, the STE test method has an accuracy of 85%, an FN rate of 12%, and an FP rate of 19% (OECD 491, 2020b). Test chemicals with vapor pressures higher than 6 kPa that fail to dissolve or form a suspension in mineral oil are outside of the application domain of the test method (OECD 491, 2020b).

"State of the Art"

None of the tests described above have been found to be capable of detecting the full range of the ocular irritation response, particularly those materials that cause an early response that resolves or "reverses" over time, that is, GHS category 2B/2A and EPA categories III and II. To address this issue, various groups have advocated for a tiered-testing strategy using multiple tests in a top-down or bottom-up approach using multiple test procedures in combination, making testing more expensive and difficult to interpret. The Integrated Approach on Testing and Assessment (IATA) described in OECD 263 (2019c) improves on the sequential testing previously described in OECD 405 (2020c) and provides decision-making guidelines. The IATA for eye hazard identification is divided into nine "modules" with three major sections that are further detailed in OECD 263. In the IATA, it is suggested that the bottom-up approach should be followed when existing information is insufficient and the weight-of-evidence (WoE) assessment results in a high probability that the test chemical does not require classification for eye hazards (GHS NC versus GHS category 1/category 2). On the other hand, a top-down approach is recommended when the WoE supports probability for GHS category 1 chemical. Hayashi et al. (2012) used a tiered-testing approach combining the data of the STE and BCOP tests for predicting eye irritation potential of chemicals concluded that for these two tests, that the accuracy of the GHS prediction was only slightly improved when the tiered approach combination was used compared to the STE test irritation rank classification alone (Hayashi et al., 2012).

The recommended maximum of three test methods in a tiered-testing strategy (Adriaens et al., 2018a) may not adequately identify reversible ocular irritants with a reasonable level of statistical certainty, as validated, high-sensitivity methods for the detection of GHS NC have a high FP rate, and validated methods for the detection of ocular corrosives (GHS category 1) have low sensitivity. This is further confounded by the fact that published accuracies of different tests use different validation procedures, making the comparisons between test methods and selection of which tests to use imprecise. Specifically, different studies use different chemicals. As demonstrated below, the mix of chemicals used to calculate accuracy and FP rates has a major impact on these statistics.

To address this, we conducted an extensive database search to identify chemicals that were used in common between the different test methods. Comparisons were then made based on these common chemicals to more precisely compare the accuracies between tests in the hope of identifying optimal tests and testing strategies.

METHODS

Database Source Selection

Search Criteria—To ensure an objective and thorough analysis, we developed a standardized search procedure to identify validation study results for the different methods to be compared: The NIH PubMed database and specific government agency websites (OECD, ICCVAM, NICEATM, ECCVAM, EURL ECVAM, and JaCVAM) were used as peer-reviewed and trusted sources. To search for test method validation results on these websites, the full name of the test method and the abbreviated form were used; Bovine Corneal Opacity and Permeability (BCOP), EpiOcular Eye Irritation (EIT), Isolated Chicken Eye (ICE), Ocular Irritrection (OI), OptiSafe (OS), and Short Time Exposure (STE). The search was performed as described below.

The NIH PubMed database (NIH, accessed July 27, 2020) was used to search for publications, the OECD website (OECD, accessed July 27,2020) was used to search for test guidelines, the U.S. NTP website (NTP, accession July 27, 2020) was used to search for ICCVAM and NICEATM reports, and the EU Science Hub website (EU, accessed July 27, 2020) was used to search for ECCVAM and EURL ECVAM reports. The following keywords were used to find the sources for each test method: "TEST METHOD NAME," "TEST METHOD NAME ABBREVIATION," "TEST METHOD NAME" + "TEST," "TEST METHOD NAME ABBREVIATION" + "TEST" and "TEST METHOD NAME" + "VALIDATION," "TEST METHOD NAME ABBREVIATION" + "VALIDATION," "TEST METHOD NAME" + "EYE," "TEST METHOD NAME ABBREVIATION" + "EYE," "TEST METHOD NAME" + "OCULAR," and "TEST METHOD NAME ABBREVIATION" + "OCULAR."

The first 100 results were analyzed from each of the search criteria above. Each result was checked to make sure the methods used were consistent with the most current test procedure and that a data set of chemicals tested with corresponding CASRN and GHS in vivo and in vitro results for each of the chemicals was included in the publication. In the case of BCOP, the sources were required to distinguish between what type of opacitometer was used. These results were used as the sources for comparisons between each test method; abbreviated results are shown in Table 1, and detailed results are shown in Supplemental Table 1.

There were a number of sources with duplicate data, as some sources compiled the data from multiple studies. For BCOP (OP-KIT), the Test Method Evaluation Report Volume 1 (ICCVAM, 2010a) contains data from the Balls et al. (1995) and Gautheron et al. (1994) studies. The BCOP method was divided into BCOP (OP-KIT) and BCOP (LLBO) to account for the different opacitometers utilized (OECD, 2020a). For ICE, the Test Method Evaluation Report Volume 2 (ICCVAM, 2010b) contains data from the Balls et al. (1995), Prinsen and Koeter (1993), and Prinsen (1996) studies. For STE, the Test Method Summary Review Document (ICCVAM-NICEATM, 2013) contains data from the Takahashi et al. (2009 and 2010) and Sakaguchi et al. (2011) studies. All data were obtained from the published literature, with the exception of isothioglycolate (CASRN 25103–09-7) results for the OptiSafe test method. The chemical was determined to be impure, and a different result was obtained when a purer lot was used. Repeated testing of 98% pure isothioglycolate

showed that the OptiSafe method accurately predicts this chemical as a GHS NC (additional details are provided in Supplemental Material I). Therefore, this chemical was not used for the OptiSafe analysis reviewed herein.

From the sources listed in Table 1, chemicals were selected for analysis if the CASRN, GHS in vivo classification, and GHS in vitro prediction were provided. Chemicals without a CASRN could not be selected for analysis because there was no way to accurately identify the same chemical tested using another test method due to variations in chemical names. Chemicals without a GHS in vivo classification were also not selected for analysis because the chemical could not be classified as a true positive (TP), FN, true negative (TN), or FP with only an in vitro classification or result. To determine if a chemical was tested in a masked study, a statement on the source that indicates which chemicals were tested "blinded", "masked," or "coded" was required. In this context, masked means that the samples to be tested are aliquoted into coded vials, and the laboratory conducting the test does not have the key to code numbers or know what is in each vial. Note: "masked" can be considered the same as "coded" or "blinded". To ensure study quality was sufficient, the following criteria were applied:

1. Only studies using the most current protocol were considered.

2. All studies were published and peer-reviewed.

3. Results were for chemicals with a CASRN and in vivo data, and could theoretically be repeated if questions arose.

4. Masked ("blinded" or "coded") versus all data comparisons were separated.

Results and Discussion

Results for the BCOP (LLBO), BCOP (OP-KIT), EpiOcular, ICE, Ocular Irritection, OptiSafe, and STE methods were compiled into a comprehensive table. Supplemental Table 2 shows the same-chemical comparisons for GHS categories 1, 2B/2A, and NC. An NC versus irritant analysis was used to determine if the chemical was a TP, FN, TN, or FP. Results for masked, blinded, or coded studies are indicated as masked (M), and not masked, not blinded, or not coded studies are indicated as NM. A dash (−) indicates there was no result provided from the sources used or there was no agreement between results to make a final classification decision. Also shown in this table is the OECD Toolbox V4.4.1 functional group for each chemical. While readers may find it bothersome to seek out supplemental tables on journal websites, given the very large number of results analyzed (363), it is not possible to put these very large tables within the published manuscript. Nonetheless, some interested parties will find these comprehensive comparison tables useful; therefore, they are included as supplemental materials. In addition, supplemental table 2 provides the reader with the capability to independently verify results and calculations, and may be useful for other applications.

Same-Chemical Accuracy Comparisons Demonstrate that the Mix of Chemicals Tested Determines Perceived Accuracy

Results for chemicals tested by two or more test methods were used to identify TP, FP, TN, and FN for chemicals in common between pairs of test methods. These values were used to determine the FP rate, FN rate, accuracy, and balanced accuracy for chemicals in common between the different test methods evaluated. Therefore, accuracy results are segregated based on if the two tests being compared assessed the same chemicals. Table 2 shows the FN rate, FP rate, overall accuracy, and balanced accuracy comparison for each pair of tests evaluated with exactly the same chemicals. By comparing the same chemicals, unknown chemical selection bias is controlled for, allowing for an evaluation of the relative accuracy of one test compared to another. Supplemental Table 3 shows the accuracy comparisons of masked (i.e., coded or "blinded") studies.

Supplemental Table 4 compares the *n* for coded study comparisons, with the *n* for all available data (resulting from both coded and not coded studies). The total number of chemicals in common, especially the number of GHS NC chemicals, are limited for coded comparisons. For many of the coded comparisons, there are fewer than 10 negatives (GHS NC) in common, and in several examples, there are just one or two negatives in common (Supplemental Table 4). The overall same chemical comparisons provide an increased number of chemicals in common for a better comparison but may still be considered a preliminary comparison because the numbers of chemicals in common are low and likely below what would be required for a robust statistical comparison. As shown in Supplemental Table 5, the balanced accuracies are similar between the masked and overall comparisons. It is interesting to note that the accuracy for some of the comparisons was slightly higher for the coded studies.

As shown in table 2, although the *n*'s are low, there are few commonly missed FNs, suggesting that the different tests likely detect different mechanisms of irritation and therefore combinations of tests likely identify most irritants. An exception is the BCOP (OP-KIT) method compared to the BCOP (LLBO). Nonetheless, as shown in Table 2, about one-third to one-half or more of the FPs are commonly missed (Table 2, commonly missed FPs) with the notable exception of the STE test.

Table 3 is a compilation of the accuracy ranges for each test presented in Table 2. the published accuracy for each test method with the range of accuracies that result from selection criteria based on the same chemicals tested by two tests (segregated based on chemicals commonly tested by the two test methods being compared in each instance). The relative FP rates are variable and different than the published FP rates, except for the STE test method, and the relative FN rates are variable and different than the published FN rates, except for the OptiSafe method. Tables 2 and 3 show a similarity in results for matched chemicals which is more pronounced than the relative accuracy of one test compared to another for the same set of chemicals. Very different accuracies for the same test depend on which chemicals are in common for a given pair comparison. This indicates that the mix of chemicals has a major impact on the perceived accuracy of the test method; in other words, a test method can be perceived as highly accurate or inaccurate, depending on the mix of chemicals evaluated.

Table 2 also demonstrates that for any given pair, there are a high percentage of commonly missed FPs. This is explored more below.

There Is a Group of Chemicals that are GHS FPs by All of the Tests Evaluated Here

To evaluate commonly missed FPs in more detail, we narrowed the number of chemicals down to those that were assessed using two or more tests, this reduced the number of chemicals from 363 to 202 chemicals. Additionally, upon closer inspection of chemical results, it was noted that the dilution for one chemical (No. 92) was not reported in two studies leading to different classifications (both irritant and nonirritant) and different results leading to TP and FP. Because of this disparity, this chemical was also removed from the data set leaving only 201 chemicals that were further analyzed.

As shown in Table 4, of the 201 chemicals tested in 2 or more tests, 121 chemicals were tested by BCOP (LLBO), 145 chemicals tested by BCOP (OP-KIT), 144 chemicals tested by EpiOcular, 67 chemicals tested by ICE, 74 chemicals tested by OI, 81 chemicals tested by OptiSafe, and 117 chemicals tested by STE. Of the chemicals tested, BCOP (LLBO), BCOP (OP-KIT), EpiOcular, and ICE had twice as many irritants as NC chemicals, while, Ocular Irritection, OptiSafe, and STE had about equal numbers for irritants and nonirritants.

Table 5 compares chemicals analyzed using two or more tests. The left hand column (out-of-sequence numbers) refers to the in-sequence numbers on the left-hand column of the Supplemental Table 2 (Supplemental Table 2 is organized by ascending CAS Registry Number and provides additional information, including functional groups and additional references). Of the 201 test chemicals, 82 were classified as NC by the GHS classification system (Table 5A), while 119 were classified as irritants (Table 5B) with 72 classified as category 1 and 47 classified as category 2 irritants. On the other hand, of the 155 EPA classifiable chemicals, 39 were classified as EPA category IV, while 55 were classified EPA category III, 23 as EPA category II, and 38 as EPA category I. Of all the tests, only 7 chemicals were tested by all 7 alternative tests, while 16 chemicals were tested by 6 tests, 30 chemicals were tested by 5 tests, 56 chemicals were tested by 4 tests, 46 chemicals were tested by 3 tests, and the remaining 46 chemicals were tested by only two tests.

Of the 82 NC chemicals, 35 were correctly classified in all tests (Table 5A). Thirty-one of these chemicals had EPA classification, and of these 31, 25 were classified as EPA category IV with 5 being classified as EPA category III and 1 classified as EPA category II. On the other hand, 17 chemicals were misclassified by all test, with 2 chemicals missed by 5 tests, 2 chemicals missed by 4 tests, 6 chemicals missed by 3 tests, and 7 chemicals missed by 2 tests. Of those chemicals with EPA classifications, 12 were classified as EPA category III with 4 classified as EPA category IV. Of the remaining NC chemicals that were FP, at least 1 test correctly classified the material; however, 13 chemicals showed correct classification in only one of the other three to five tests. EPA classification for these chemicals also showed a greater number of EPA category III chemicals (7) compared to EPA category IV (3). Of these 13 missed chemicals, no one single test correctly classified them, with Ocular Irritection correctly classifying one chemical, EpiOcular correctly classifying 4 chemicals and STE correctly classifying 8 chemicals. Table 5A demonstrates there are

commonly misclassified chemicals, and many of the commonly overclassified chemicals are EPA category III/GHS NC.

Based on an extensive review of the literature, this paper establishes a historical database of test results of BCOP (LLBO), BCOP (OP-KIT), EpiOcular, ICE, Ocular Irritection, OptiSafe, and STE that provides GHS classifications and functional group identifications for which comparative accuracies and empirical tiered-testing strategies can be analyzed. Importantly, using this database, we have identified a group of chemicals that provide FP results in all reported testing strategies (Lower 1/3 of Table 5A). Specifically, there is a group of chemicals that are misclassified as FP by all in vitro eye irritation tests that were identified. While the STE method alone did show a lower FP rate for some of these chemicals, albeit also showing the highest FN, there was a high percent of FPs that are missed between test methods. When some chemicals are commonly missed by all tests, inclusion of these chemicals in a validation study will have an adverse impact on perceived test method accuracy.

There are several potential explanations for why all identified alternative tests misclassify this set of chemicals. First, we did identify that as a group, these chemicals contained specific functional groups in common, specifically it was identified that the amine, aryl and aryl halide, ether, ketone, and alkene were disproportionally overclassified by every test except for the STE test. However, it should be noted that these functional groups were not unique to the set of missed chemicals, and therefore cannot be used as markers for this set of chemicals. Further comparisons of functional groups and other chemical properties with misprediction rates and specific ocular effects may provide insights into the chemistry and biological processes of these chemicals. However, the n is low, and these preliminary results are best suited for hypotheses and as the bases of prospective studies, for example, a guess that other chemicals with these functional groups will also be FPs for a given test method. If the database were expanded in the future, a program might be developed that would allow the input of a chemical structure to yield a suggestion of which in vitro test would have the highest accuracy for the associated functional groups.

Another possible explanation for the mispredictions for this set of chemicals and their functional groups may be the variability of the in vivo database. To the extent that the in vivo data for a specific material do not reflect a risk to humans, the in vivo data are of little value for risk assessment and distracts from a meaningful validation of nonanimal test methods. Consistently mispredicted chemicals should be identified and flagged, and then additional studies of human exposure data and chemical and toxicological properties performed to allow stakeholders to develop consensus opinions on the appropriateness to include these chemicals as gold standards for validation studies.

Finally, common mispredictions may be due to the inability of all the tests evaluated to detect chemical irritants that show reversibility. Greater than 80% of the consistently overpredicted NC chemicals with EPA classifications, were classified as EPA category III (Table 5A). Therefore, it appears that the test methods discussed here do not differentiate between damage at 24 hours and damage at 3 days, resulting in a high GHS FP rates for most tests when EPA category III/GHS NC classified chemicals are included in the

validation. This finding suggests that there is still a need for further alternative test development based on detection of reversibility, if there is a desire to reduce the number of FPs. The need for alternative tests to address mechanisms of reversibility is also supported by the evaluation of the effects of these persistently missed chemicals on tiered-testing strategies.

Tiered Testing with Two Tests Provides Little or No Advantage over a Single Test

Same chemical matched results for two tests can be used to evaluate sequential testing strategies for the same two tests. When the FP rate is high, as is the case for nonanimal eye irritation tests, a common bottom up, tiered-testing strategy is to retest positives using a second test. This strategy theoretically reduces the FP rate (because FPs have a second chance to be classified as TNs) but may increase the FN rate (because TPs have a second chance to misclassified as FNs). In Table 6, the theoretical and actual accuracies for the various combinations of testing strategies are provided when positives are then evaluated by a second test. For the retesting of positives, it does not matter which test is first. The theoretical combined FN rate is always {100% − [(100% − FN% of test 1) • (100% − FN% of test 2)]}, regardless of which test is first. For example, as shown in 6A (EpiOcular and ICE), {1 − [(1 − 0.029) • (1 − 0.059)] = 8.7%}. FNs are typically increased by the second test. On the other hand, FP values are typically reduced by the second test. For example, if EpiOcular is the first test (FP rate = 44.4%), then positives are retested by ICE (FP rate = 66.7%), the theoretical FP rate of the combined tests is 0.444 • 0.667 = 29.6%. Likewise, if ICE is first and EpiOcular is second, the combined FP rate is the same (0.667 • 0.444 = 29.6%).

However, results shown in Table 6 demonstrate that the bottom-up strategy of retesting positives typically results in an increased FN rate. If FPs are commonly misclassified, the actual improvement in specificity is not as good as the theoretical improvement. As shown in Table 6, in some cases, the actual results from just one test are more accurate than retesting positives with a second test (because the second test contributes no or a limited actual reduction of FPs but does increase the FN rate). Because the same chemicals are commonly overclassified, retesting positives by a second test method results in a lower-than-expected improvement of the FP rate and typically increases the FN rate (Table 6).

In summary, as demonstrated by same-chemical comparisons, perceived accuracy of one test method compared to another is highly dependent on the selection of chemicals. In addition, the tests evaluated herein miss many of the same chemicals and functional groups in the substances analyzed; in these cases, a second test in a tier will not add statistical power to a bottom-up analysis. It is suggested that chemicals misclassified using these test methods be further studied. Specifically studying biological endpoints associated with mispredictions would allow specific mechanisms related to mispredictions to be identified and enable the development of more predictive and biologically relevant nonanimal tests.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

REFERENCES

Adriaens E, Verstraelen S, Alépée N, Kandarova H, Drzewiecka A, Gruszka K, Guest R, Willoughby JA Sr., Van Rompay AR. (2018a). "CON4EI: Development of testing strategies for hazard identification and labelling for serious eye damage and eye irritation of chemicals". Toxicol In Vitro. 49:99–115. doi: 10.1016/j.tiv.2017.09.008 [PubMed: 28964898]

Adriaens E, Willoughby JA Sr., Meyer BR., Blakeman LC., Alépée N, Fochtman P, Guest R, Kandarova H, Verstraelen S, Van Rompay AR. (2018b). "CON4EI: Short Time Exposure (STE) test method for hazard identification and labelling of eye irritating chemicals". Toxicol In Vitro. 49:65–76. doi: 10.1016/j.tiv.2017.08.002. [PubMed: 28801046]

Alépée N, Bessou-Touya S, Cotovio J, de Smedt A, de Wever B, Faller C, Jones P, Le Varlet B, Marrec-Fairley M, Pfannenbecker U, Tailhardat M, van Goethem F, McNamee P (2013). "Cosmetics Europe multi-laboratory pre-validation of the SkinEthic™ reconstituted human corneal epithelium test method for the prediction of eye irritation". Toxicol In Vitro. 27(5):1476–88. doi: 10.1016/j.tiv.2013.02.009. [PubMed: 23524228]

Alépée N, Adriaens E, Abo T, Bagley D, Desprez B, Hibatallah J, Mewes K, Pfannenbecker U, Sala À, Van Rompay AR., Verstraelen S, McNamee P (2019). "Development of a defined approach for eye irritation or serious eye damage for liquids, neat and in dilution, based on cosmetics Europe analysis of in vitro STE and BCOP test methods". Toxicol In Vitro. 57:154–163. doi: 10.1016/j.tiv.2019.02.019. [PubMed: 30817952]

ALTEX. (2019). "Successes with State Cosmetics Legislation builds momentum for US federal nan. ALTEX: Alternatives to Animal Experimentation, Corners. 36(3): 497. Available at: https://www.altex.org/index.php/altex/article/view/1312/1314

Balls M, Botham PA, Bruner LH, Spielmann (1995). "The EC/HO international validation study on alternatives to the Draize eye irritation test." Toxicol. In Vitro 9, 871–929. [PubMed: 20650169]

Barroso J, Pfannenbecker U, Adriaens E, Alepee N, Cluzel M, Smedt AD., Hibatallah J, Klaric M, Mewes KR., Millet M, Templier M, McNamee P (2017). "Cosmetics Europe compilation of historical serious eye damage/eye irritation in vivo data analysed by drivers of classification to support the selection of chemicals for development and evaluation of alternative methods/strategies: the Draize eye test Reference Database (DRD)". Arch Toxicol. 91(2):521–547. doi:10.1007/s00204-016-1679-x [PubMed: 26997338]

Choksi N, Lebrun S, Nguyen M, Daniel A, DeGeorge G, Willoughby J, Layton A, Lowther D, Merrill J, Matheson J, Barroso K, Yozzo K, Casey W, Allen D (2020). "Validation of the OptiSafe™ eye irritation test". Cutan Ocul Toxicol. 39(3):180–192. doi:10.1080/15569527.2020.1787431 [PubMed: 32586141]

Doucet O, Lanvin M, Thillou C Linossier C, Pupat C, Merlin B, Zastrow L (2006). "Reconstituted human corneal epithelium: a new alternative to the Draize eye test for the assessment of the eye irritation potential of chemicals and cosmetic products". Toxicol In Vitro. 20(4):499–512. doi:10.1016/j.tiv.2005.09.005 [PubMed: 16243479]

ECETOC. (1992). "Technical Report No. 48, Eye Irritation Reference Chemicals Data Bank". European Chemical Industry Ecology and Toxicology Centre, Brussels.

ECETOC. (1998). "Technical Report No. 48(2), Eye Irritation Reference Chemicals Data Bank (Second Edition)".European Chemical Industry Ecology and Toxicology Centre, Brussels. Available at: ecetoc.org/wp-content/uploads/2014/08/ECETOC-TR-0481.pdf

Eskes C, Hoffmann S, Facchini D, Ulmer R, Wang A, Flego M, Vassallo M, Bufo M, van Vliet E, d'Abrosca F, Wilt N (2014). "Validation study on the Ocular Irritection assay for eye irritation testing". Toxicol In Vitro. 28(5):1046–65. doi: 10.1016/j.tiv.2014.02.009. [PubMed: 24637247]

EU. Accessed July 27, 2020. "EU Science Hub: The European Commission's science and knowledge service". Available at: https://ec.europa.eu/jrc/en

EURL ECVAM. (2014). "The EURL ECVAM – Cosmetics Europe prospective validation study of Reconstructed human Cornea-like Epithelium (RhCE)-based test methods for identifying chemicals not requiring classification and labeling for serious eye damage/eye irritation". European Commission. Available at: https://publications.jrc.ec.europa.eu/repository/bitstream/JRC100280/lana28125enn.pdf

European Community. (2009). "Regulation (EC) No 1223/2009 of the European Parliament and of the Council of 30 November 2009 on cosmetic products". Official Journal of the European Union. Available at: https://ec.europa.eu/health/sites/health/files/endocrine_disruptors/docs/cosmetic_1223_2009_regulation_en.pdf

Gautheron P, Giroux J, Cottin M, Audegond L, Morilla A, Mayordomo-Blanco L, Tortajada A, Haynes G, Vericat JA., Pirovano R, Gillio Tos E, Hagemann C, Vanparys P, Deknudt G, Jacobs G, Prinsen M, Kalweit S, Spielmann H (1994). "Interlaboratory assessment of the bovine corneal opacity and permeability (BCOP) assay". Toxicol In Vitro. 8(3):381–92. doi:10.1016/0887-2333(94)90159-7. [PubMed: 20692929]

Grum T (2019). "Global ban on animal testing: where are we in 2019?". Cosmetics Design-Europe. Available at: https://www.cosmeticsdesign-europe.com/Article/2019/03/05/Global-ban-on-animal-testing-where-are-we-in-2019?utm_source=copyright&utm_medium=OnSite&utm_campaign=copyright

Hayashi K, Mori T, Abo T, Koike M, Takahashi Y, Sakaguchi H, Nishiyama N (2012). "A tiered approach combining the short time exposure (STE) test and the bovine corneal opacity and permeability (BCOP) assay for predicting eye irritation potential of chemicals". J Toxicol Sci. 37(2):269–80. doi.org/10.2131/jts.37.269 [PubMed: 22467017]

Hayashi K, Abo T, Nukada Y, Sakaguchi H (2013). "Definition of the applicability domain of the Short Time Exposure (STE) test for predicting the eye irritation of chemicals". Altern Lab Anim. 41(2):157–71. doi: 10.1177/026119291304100205. [PubMed: 23781933]

H.R.5141. 2019. Humane Cosmetics Act 116th Congress. Rep. Beyer, Donald S., Jr. [D-VA-8] (Introduced 11/18/2019) [Online] Available at: https://www.congress.gov/bill/116th-congress/house-bill/5141/text?r=5&s=1

ICCVAM. (2010a). "ICCVAM Recommended BCOP Test Method Protocol. In: ICCVAM Test Method Evaluation Report – Current Validation Status of in vitro Test Methods Proposed for Identifying Eye Injury Hazard Potential of Chemicals and Products. Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)". NIH Publication No.: 10–7553A. Available: [https://ntp.niehs.nih.gov/iccvam/docs/ocutox_docs/invitro-2010/tmer-vol1.pdf].

ICCVAM. (2010b). "ICCVAM Test Method Evaluation Report: Current Validation Status of In Vitro Test Methods Proposed for Identifying Eye Injury Hazard Potential of Chemicals and Products (Volume 2). Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM)". NIH Publication No.: 10–7553. Available at: https://ntp.niehs.nih.gov/iccvam/docs/ocutox_docs/invitro-2010/tmer-vol2.pdf

ICCVAM-NICEATM. (2013). "Short Time Exposure (STE) Test Method Summary Review Document", National Toxicology Program. Available at: http://www.ntp.niehs.nih.gov/iccvam/docs/ocutox_docs/STE-SRD-NICEATM-508.pdf.

Jung KM., Lee SH., Ryu YH., Jang WH., Jung HS., Han JH., Seok SH., Park JH., Son Y, Park YH., Lim KM. (2011). "A new 3D reconstituted human corneal epithelium model as an alternative method for the eye irritation test". Toxicol In Vitro. 25(1):403–410. doi:10.1016/j.tiv.2010.10.019 [PubMed: 21040779]

Kaluzhny Y, Kandárová H, Hayden P, Kubilus J, d'Argembeau-Thornton L, Klausner M (2011). "Development of the EpiOcular(TM) eye irritation test for hazard identification and labelling of eye irritating chemicals in response to the requirements of the EU cosmetics directive and REACH legislation". Altern Lab Anim. 9;39(4):339–64. doi: 10.1177/026119291103900409. [PubMed: 21942547]

Kandarova H, Letasiova S, Adriaens E, Guest R, Willoughby JA Sr., Drzewiecka A, Gruszka K, Alépée N, Verstraelen S, Van Rompay AR. (2017). "CON4EI: EpiOcular™ Eye Irritation Test (EpiOcular™ EIT) for hazard identification and labelling of eye irritating chemicals". Toxicol In Vitro. 49:21–33. doi: 10.1016/j.tiv.2017.07.002. [PubMed: 28697962]

Katoh M, Hamajima F, Ogasawara T, Hata K (2013). "Establishment of a new in vitro test method for evaluation of eye irritancy using a reconstructed human corneal epithelial model, LabCyte CORNEA-MODEL". Toxicol In Vitro. 27(8):2184–92. doi: 10.1016/j.tiv.2013.08.008. [PubMed: 23999412]

Kojima H, Hayashi K, Sakaguchi H, Omori T, Otoizumi T, Sozu T, Kuwahara H, Hayashi T, Sakaguchi M, Toyoda A, Goto H, Watanabe S, Ahiko K, Nakamura T, Morimoto T (2013). "Second-phase validation study of short time exposure test for assessment of eye irritation potency of chemicals". Toxicol In Vitro. 2013 9;27(6):1855–69. [PubMed: 23747838]

Kolle SN., Kandárová H, Wareing B, van Ravenzwaay B, Landsiedel R (2011). "In-house validation of the EpiOcular(TM) eye irritation test and its combination with the bovine corneal opacity and permeability test for the assessment of ocular irritation". Altern Lab Anim. 39(4):365–87. doi: 10.1177/026119291103900410. [PubMed: 21942548]

Lebrun S, Xie Y, Chavez S, Chan R, Jester JV (2019). "An in vitro depth of injury prediction model for a histopathologic classification of EPA and GHS eye irritants." Toxicol In Vitro. 61:104628. doi: 10.1016/j.tiv.2019.104628 [PubMed: 31419508]

Luechtefeld T, Maertens A, Russo DP., Rovida C, Zhu H, Hartung T (2016a). "Analysis of Draize eye irritation testing and its prediction by mining publicly available 2008–2014 REACH data". ALTEX: Alternatives to Animal Experimentation. 33(2):123–134. doi:10.14573/altex.1510053 [PubMed: 26863293]

Luechtefeld T, Maertens A, Russo DP., Rovida C, Zhu H, Hartung T (2016). "Analysis of publicly available skin sensitization data from REACH registrations 2008–2014". ALTEX: Alternatives to Animal Experimentation. 33(2):135–148. doi:10.14573/altex.151005. [PubMed: 26863411]

Lim SE., Ha SJ., Jang WH., Jung KM., Jung MS., Yeo KW., Kim JS., Jeong TC., Kang MJ., Kim SY., Lee SH., Ko KY., Kim TS., Park KS., Bae S, Lim KM. (2019). "Me-too validation study for in vitro eye irritation test with 3D-reconstructed human cornea epithelium, MCTT HCE™". Toxicol In Vitro. 55:173–184. doi: 10.1016/j.tiv.2018.12.003. [PubMed: 30572010]

NIH. Accessed July 27, 2020. "National Library of Medicine National Center for Biotechnology Information, PubMed.gov". Available at: https://pubmed.ncbi.nlm.nih.gov/

NTP. Accessed July 27, 2020. "National Toxicology Program U.S. Department of Health and Human Services". Available at: https://ntp.niehs.nih.gov/

OECD. (1998). "OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1". Available at: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/mc/chem(98)17&doclanguage=en

OECD. (2018). "Test No. 438: Isolated Chicken Eye Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage". OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/9789264203860-en.

OECD. (2019a). "Test No. 492: Reconstructed human Cornea-like Epithelium (RhCE) test method for identifying chemicals not requiring classification and labelling for eye irritation or serious eye damage". OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/9789264242548-en.

OECD. (2019b). "Test No. 496: in vitro Macromolecular Test Method for Identifying Chemicals Inducing Serious Eye Damage and Chemicals not Requiring Classification for Eye Irritation or Serious Eye Damage. OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/970e5cd9-en.

OECD. (2019c). "Guidance Document No. 263 on Integrated Approaches to Testing and Assessment (IATA) for Serious Eye Damage and Eye Irritation Series on Testing and Assessment Number 263 (Second Edition)." http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2017)15/REV1&docLanguage=En

OECD. (2020a). "Test No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for

Eye Irritation or Serious Eye Damage". OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/9789264203846-en

OECD. (2020b). "Test No. 491: Short Time Exposure In Vitro Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage". OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/9789264242432-en.

OECD. (2020c). "Test No. 405: Acute Eye Irritation/Corrosion" OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, 10.1787/9789264185333-en.

OECD. Accessed July 27, 2020. "OECD Better Policies for Better Lives". Available at: https://www.oecd.org/

Pfannenbecker U, Bessou-Touya S, Faller C, Harbell J, Jacob T, Raabe H, Tailhardat M, Alépée N, De Smedt A, De Wever B, Jones P, Kaluzhny Y, Le Varlet B, McNamee P, Marrec-Fairley M, Van Goethem F (2013). "Cosmetics Europe multi-laboratory pre-validation of the EpiOcular™ reconstituted human tissue test method for the prediction of eye irritation". Toxicol In Vitro. 3;27(2):619–26. doi: 10.1016/j.tiv.2012.11.007. [PubMed: 23159500]

Prinsen MK. (1996). "The chicken enucleated eye test (CEET): a practical (pre)screen for the assessment of eye irritation/corrosion potential of test materials". Food Chem Toxicol. 34(3):291–6. doi:10.1016/0278-6915(95)00115-8. [PubMed: 8621111]

Prinsen MK., Koëter HB. (1993). "Justification of the enucleated eye test with eyes of slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits". Food Chem Toxicol. 31(1):69–76. doi:10.1016/0278-6915(93)90182-x. [PubMed: 8444390]

Sakaguchi H, Ota N, Omori T, Kuwahara H, Sozu T, Takagi Y, Takahashi Y, Tanigawa K, Nakanishi M, Nakamura T, Morimoto T, Wakuri S, Okamoto Y, Sakaguchi M, Hayashi T, Hanji T, Watanabe S (2011). "Validation study of the Short Time Exposure (STE) test to assess the eye irritation potential of chemicals". Toxicol In Vitro. 25(4):796–809. doi:10.1016/j.tiv.2011.01.015. [PubMed: 21291995]

Takahashi Y, Hayashi T, Watanabe S, Hayashi K, Koike M, Aisawa N, Ebata S, Sakaguchi H, Nakamura T, Kuwahara H, Nishiyama N (2009). "Inter-laboratory study of short time exposure (STE) test for predicting eye irritation potential of chemicals and correspondence to globally harmonized system (GHS) classification". J Toxicol Sci. 34(6):611–26. doi:10.2131/jts.34.611. [PubMed: 19952497]

Takahashi Y, Hayashi T, Koike M, Sakaguchi H, Kuwahara H, Nishiyama N (2010). "An interlaboratory study of the short time exposure (STE) test using SIRC cells for predicting eye irritation potential". Cutan Ocul Toxicol. 29(2):77–90. doi:10.3109/15569521003587327. [PubMed: 20178401]

Takahashi Y, Hayashi K, Abo T, Koike M, Sakaguchi H, Nishiyama N (2011). "The Short Time Exposure (STE) test for predicting eye irritation potential: intra-laboratory reproducibility and correspondence to globally harmonized system (GHS) and EU eye irritation classification for 109 chemicals". Toxicol In Vitro. 25(7):1425–34. doi: 10.1016/j.tiv.2011.04.012. [PubMed: 21513790]

Verstraelen S, Jacobs A, De Wever B, Vanparys P (2013). "Improvement of the Bovine Corneal Opacity and Permeability (BCOP) assay as an in vitro alternative to the Draize rabbit eye irritation test". Toxicol In Vitro. 27(4):1298–311. doi: 10.1016/j.tiv.2013.02.018. [PubMed: 23501624]

Verstraelen S, Maglennon G, Hollanders K, Boonen F, Adriaens E, Alépée N, Drzewiecka A, Gruszka K, Kandarova H, Willoughby JA Sr., Guest R, Schofield J, Van Rompay AR. (2017). "CON4EI: Bovine Corneal Opacity and Permeability (BCOP) test for hazard identification and labelling of eye irritating chemicals". Toxicol In Vitro. 44:122–133. doi: 10.1016/j.tiv.2017.06.028. [PubMed: 28673559]

Weil CS., Scala RA. (1971). "Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation test". Toxicol Appl Pharmacol. 19(2):276–360. doi: 10.1016/0041-008X(71)90112-8. [PubMed: 5570968]

Yang H, Kim DE., Jang WH., An S, Cho SA., Jung MS., Lee JE., Yeo KW., Koh SB., Jeong TC., Kang MJ., Chun YJ., Lee SH., Lim KM., Bae S (2017). "Prevalidation trial for a novel in vitro eye irritation test using the reconstructed human cornea-like epithelial model, MCTT HCE™". Toxicol In Vitro. 39:58–67. doi: 10.1016/j.tiv.2016.11.010. [PubMed: 27876558]

York M, Steiling W (1998). "A critical review of the assessment of eye irritation potential using the Draize rabbit eye test". J Appl Toxicol. 18(4):233–240. doi:10.1002/(sici)1099-1263(199807/08)18:4<233::aid-jat496>3.0.co;2-y [PubMed: 9719422]

Comparison of In Vitro Eye Irritation Tests: Bovine Corneal Opacity and Permeability, EpiOcular™, Isolated Chicken Eye, Ocular Irritection™, OptiSafe™, and Short Time Exposure.

Same-chemical comparisons demonstrate that the perceived accuracy of one test compared to another is dependent on the specific chemicals used for the validation study.

Some chemicals and functional groups appear to be broadly misclassified by all test methods.

To the extent that all tests make the same misclassifications, the advantage of using a tiered-testing strategy is reduced.

Table 1.

Sources Identified From the Database Search Criteria for the BCOP (LLBO), BCOP (OP-KIT), EpiOcular, ICE, Ocular Irritection, OptiSafe, and STE Test Methods.

Test Method	Sources
BCOP (LLBO)	Adriaens et al., 2018a [1], Verstraelen et al., 2013 [4], Verstraelen et al., 2017 [5], OECD, 2020a [6], Alepee et al., 2019 [8]
BCOP (OP-KIT)	Adriaens et al., 2018a [1], Balls et al., 1995 [2] (also referenced in ICCVAM, 2010a [7]), Gautherton et al., 1994 [3] (also referenced in ICCVAM, 2010a [7]), Verstraelen et al., 2013 [4], Verstraelen et al., 2017 [5], OECD, 2020a [6], ICCVAM, 2010a [7]
EpiOcular	Adriaens et al., 2018a [1], Kandarova et al., 2017 [9], Kolle et al., 2011 [10], Pfannenbecker et al., 2013 [11], OECD, 2019a [12], EURL ECVAM, 2014 [13]
ICE	Balls et al., 1995 [2] (also referenced in ICCVAM, 2010b [17]), Prinsen and Koeter, 1993 [14] (also referenced in ICCVAM, 2010b [17]), Prinsen, 1996 [15] (also referenced in ICCVAM, 2010b [17]), OECD, 2018 [16], ICCVAM, 2010b [17]
Ocular Irritection	Eskes et al., 2014 [18], OECD, 2019b [19]
OptiSafe	Choksi et al., 2020 (20)
STE	Adriaens et al., 2018a [1], Adriaens et al., 2018b [21], Hayashi et al., 2012 [22], Hayashi et al., 2013 [23], Kojima et al., 2013 [24], Takahashi et al., 2009 [25] (also referenced in ICCVAM-NICEATM, 2013 [30]), Takahashi et al., 2010 [26] (also referenced in ICCVAM-NICEATM, 2013 [30]), Takahashi et al., 2011 [27], Sakaguchi et al., 2011 [28] (also referenced in ICCVAM-NICEATM, 2013 [30]), OECD, 2020b [29], ICCVAM-NICEATM, 2013 [30]

Table 1. Full references are listed in the reference section. Numbers in brackets corespond to the numbers in Supplemental Table 2 to indicate which source was referenced for Supplemental Table 2. BCOP = Bovine Corneal Opacity and Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; STE = Short Time Exposure; ICCVAM = Interagency Coordinating Committee on the Validation of Alternative Methods; OECD = Organization for Economic Co-Operation and Development; EURL ECVAM = European Union Reference Laboratory for alternatives to animal testing; NICEATM = National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods; NIH = National Institutes of Health. Additional details are available in Supplemental Table 1.

Table 2.

Overall Accuracy Comparison of the Same Chemicals between BCOP (LLBO), BCOP (OP-KIT), EpiOcular, ICE, Ocular Irritection, OptiSafe, and STE

Test Comparisons	n	False Negative		False Positive		Accuracy (%)	Balanced Accuracy(%)
		% (missed/total)	Commonly Missed % (same/total)	% (missed/total)	Commonly Missed % (same/total)		
BCOP (LLBO)	112	7.2 (6/83)	45.5 (5/11)	58.6 (17/29)	88.2 (15/17)	79.5 (89/112)	67.1
BCOP (OPKIT)		12.0 (10/83)		51.7 (15/29)		77.7 (87/112)	68.1
BCOP (LLBO)	104	6.3 (5/79)	0.0 (0/7)	48.0 (12/25)	38.5 (5/13)	83.7 (87/104)	72.8
EpiOcular		2.5 (2/79)		24.0 (6/25)		92.3 (96/104)	86.7
BCOP (LLBO)	37	7.1 (2/28)	0.0 (0/4)	66.7 (6/9)	50.0 (4/8)	78.4 (29/37)	63.1
ICE		7.1 (2/28)		66.7 (6/9)		78.4 (29/37)	63.1
BCOP (LLBO)	32	9.5 (2/21)	0.0 (0/3)	45.5 (5/11)	66.7 (4/6)	78.1 (25/32)	72.5
Ocular Irritection		4.8 (1/21)		45.5 (5/11)		81.3 (26/32)	74.9
BCOP (LLBO)	41	3.6 (1/28)	0.0 (0/1)	30.8 (4/13)	75.0 (3/4)	87.8 (36/41)	82.8
OptiSafe		0.0 (0/28)		23.1 (3/13)		92.7 (38/41)	88.5
BCOP (OP-KIT)	111	10.2 (9/88)	10.0 (1/10)	47.8 (11/23)	63.6 (7/11)	82.0 (91/111)	71.0
EpiOcular		2.3 (2/88)		30.4 (7/23)		91.9 (102/111)	83.6
BCOP (OP-KIT)	61	6.8 (3/44)	0.0 (0/5)	64.7 (11/17)	66.7 (10/15)	77.0 (47/61)	64.2
ICE		4.5 (2/44)		82.4 (14/17)		73.8 (45/61)	56.6
BCOP (OP-KIT)	44	9.1 (3/33)	0.0 (0/5)	45.5 (5/11)	66.7 (4/6)	81.8 (36/44)	72.7
Ocular Irritection		6.1 (2/33)		45.5 (5/11)		84.1 (37/44)	74.2
BCOP (OP-KIT)	43	8.8 (3/34)	0.0 (0/3)	33.3 (3/9)	100.0 (3/3)	86.0 (37/43)	78.9
OptiSafe		0.0 (0/34)		33.3 (3/9)		93.0 (40/43)	83.3
EpiOcular	43	2.9 (1/34)	50.0 (1/2)	44.4 (4/9)	42.9 (3/7)	88.4 (38/43)	76.3
ICE		5.9 (2/34)		66.7 (6/9)		81.4 (35/43)	63.7
EpiOcular	58	3.3 (1/30)	0.0 (0/3)	35.7 (10/28)	90.9 (10/11)	81.0 (47/58)	80.5
Ocular Irritection		6.7 (2/30)		39.3 (11/28)		77.6 (45/58)	77.0
EpiOcular	56	2.9 (1/35)	0.0 (0/1)	33.3 (7/21)	55.6 (5/9)	85.7 (48/56)	81.9
OptiSafe		0.0 (0/35)		33.3 (7/21)		87.5 (49/56)	83.3
ICE	28	4.8 (1/21)	0.0 (0/1)	100.0 (7/7)	42.9 (3/7)	71.4 (20/28)	47.6
Ocular Irritection		0.0 (0/21)		42.9 (3/7)		89.3 (25/28)	78.6
ICE	24	10.0 (2/20)	0.0 (0/2)	75.0 (3/4)	66.7 (2/3)	79.2 (19/24)	57.5
OptiSafe		0.0 (0/20)		50.0 (2/4)		91.7 (22/24)	75.0

Test Comparisons	n	False Negative		False Positive		Accuracy (%)	Balanced Accuracy(%)
		% (missed/total)	Commonly Missed % (same/total)	% (missed/total)	Commonly Missed % (same/total)		
Ocular Irritection OptiSafe	46	8.7 (2/23) 0.0 (0/23)	0.0 (0/2)	47.8 (11/23) 43.5 (10/23)	61.5 (8/13)	71.7 (33/46) 78.3 (36/46)	71.7 78.3
STE BCOP (LLBO)	69	*20.8 (10/48) *2.1 (1/47)	10.0 (1/10)	*0.0 (0/21) *40.9 (9/22)	0.0 (0/9)	85.5 (59/69) 85.5 (59/69)	89.6 78.5
STE BCOP (OP-KIT)	84	*19.7 (12/61) *6.7 (4/60)	14.3 (2/14)	*17.4 (4/23) *50.0 (12/24)	23.1 (3/13)	81.0 (68/84) 81.0 (68/84)	81.5 71.7
STE EpiOcular	80	23.5 (12/51) 3.9 (2/51)	7.7 (1/13)	10.3 (3/29) 34.5 (10/29)	18.2 (2/11)	81.3 (65/80) 85.0 (68/80)	83.1 80.8
STE ICE	41	25.0 (7/28) 3.6 (1/28)	14.3 (1/7)	15.4 (2/13) 69.2 (9/13)	22.2 (2/9)	78.0 (32/41) 75.6 (31/41)	79.8 63.6
STE Ocular Irritection	40	33.3 (7/21) 0.0 (0/21)	0.0 (0/7)	10.5 (2/19) 42.1 (8/19)	11.1 (1/9)	77.5 (31/40) 80.0 (32/40)	78.1 78.9
STE OptiSafe	56	31.0 (9/29) 0.0 (0/29)	0.0 (0/9)	7.4 (2/27) 25.9 (7/27)	28.6 (2/7)	80.4 (45/56) 87.5 (49/56)	80.8 87.0

Table 2. Overall accuracy comparison of chemicals in common.

n is the number of chemicals in common between the test methods compared.

False negative (FN) is the number of positives misclassified as negatives [((FN) / (FN + TP)) • 100].

False positive (FP) is the number of negatives misclassified as positives [((FP) / (FP + TN)) • 100].

% (Missed/total) is the misclassification rate for chemicals in common. Commonly missed is the percent overlapping FNs or FPs.

Accuracy is the total correct predictions divided by the n in common. [((TP + TN) / (TP + TN + FP + FN)) • 100].

Balanced accuracy equally weights positives and negatives [((Sensitivity + Specificity) / 2) • 100].

BCOP = Bovine Corneal Opacity and Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; STE = Short Time Exposure. TP = True Positive; TN = True Negative; Sensitivity = [(TP) / (TP + FN)]; Specificity = [(TN) / (TN + FP)].

* = has differing total FN and FP values in common due to differing in vivo classification of cetylpyridinium bromide. The in vivo classification of cetylpyridinium bromide for LLBO and OP-KIT was not classified (NC) and the classification for STE was category 1.

Table 3.

Published Accuracy Compared to Relative Accuracy

Test Method	Published Accuracy	Relative Accuracy	Published FP rate	Relative FP rate	Published FN rate	Relative FN rate	Reference for Published
BCOP (LLBO)	83%	78.1 – 87.8%	45%	30.8 – 66.7%	6%	2.1 – 9.5%	OECD, 2020a
BCOP (OP-KIT)	69%	77.0 – 86.0%	69%	33.3 – 64.7%	0%	6.7 – 12.0%	OECD, 2020a
EpiOcular	80%	81.0 – 92.2%	37%	24.0 – 44.4%	4%	2.3 – 3.9%	OECD, 2019a
ICE	88%	71.4 – 81.4%	24%	66.7 –100.0%	3%	3.6 – 10.0%	OECD, 2018
OI	75%	71.7 – 89.3%	41%	39.3 – 47.8%	9%	0.0 – 8.7%	OECD, 2019b
OS	79–80%	78.3 – 93.0%	40%	23.1 – 50.0%	0%	0.0 – 0.0%	Choksi et al., 2020
STE	85%	77.5 – 85.5%	19%	0.0 – 17.4%	12%	19.7 – 33.3%	OECD, 2020b

Table 3. Comparison of published statistics versus relative statistics. BCOP = Bovine Corneal Opacity & Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; OI = Ocular Irritection; OS = OptiSafe; STE = Short Time Exposure; FP = false positive; FN = false negative.

Table 4.

Classification of Chemicals Analyzed by Two or More Alternative Tests.

Representative Chemicals	Alternative Tests							
	BCOP (LLBO)	BCOP (OP-KIT)	EpiOcular	ICE	OI	OS	STE	
Total Number	121	145	144	67	74	81	117
Non Classified	35	40	50	18	37	37	51
Irritant Classified	86	105	94	49	37	44	66
True Negatives	16	16	33	4	22	24	46
True Positives	80	95	92	46	34	44	*53
False Negatives	6	10	2	3	3	0	13
False Positives	*19	*24	17	14	15	13	5

Table 4. Classification of chemicals analyzed by two or more alternative tests. BCOP = Bovine Corneal Opacity and Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; OI = Ocular Irritection, OS = OptiSafe, STE = Short Time Exposure.

*
= has one less FP for BCOP (LLBO) and BCOP (OP-KIT) and one less TP for STE due to differing in vivo classification of cetylpyridinium bromide. The in vivo classification of cetylpyridinium bromide for LLBO and OP-KIT was not classified (NC) and the classification for STE was category 1.

Table 5A.

Identification of Non-Classified Chemicals.

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	TN	FP	%
41	n-Hexane	110-54-3	IV	TN	TN	TN	TN		TN	TN	6	6	0	100%
97	Potassium tetrafluoroborate	14075-53-7	IV	TN	TN	TN		TN	TN		5	5	0	100%
246	Dioctyl ether	629-82-3	IV	TN	TN	TN		TN	TN		5	5	0	100%
52	1-Bromo-octane	111-83-1	IV	TN		TN		TN	TN	TN	5	5	0	100%
207	Propylene glycol	57-55-6	IV	TN	TN	TN			TN	TN	5	5	0	100%
48	1-Bromohexane	111-25-1	IV			TN		TN	TN	TN	4	4	0	100%
89	sec-Butylbenzene	135-98-8	IV			TN		TN	TN	TN	4	4	0	100%
245	Dipropyl disulphide	629-19-6	IV			TN		TN	TN	TN	4	4	0	100%
292	2-Ethylhexyl thioglycolate	7659-86-1	IV			TN		TN	TN	TN	4	4	0	100%
238	3-Methoxy-1,2-propanediol	623-39-2	IV	TN		TN			TN	TN	4	4	0	100%
336	Tween 80	9005-65-6	IV	TN	TN	TN				TN	4	4	0	100%
133	2-Ethylhexyl p–dimethylamino benzoate	21245-02-3	IV	TN	TN	TN				TN	4	4	0	100%
167	1-Ethyl-3-methylimidazolium ethyl sulphate	342573-75-5		TN		TN				TN	3	3	0	100%
143	EDTA, di-potassium salt	25102-12-9		TN	TN					TN	3	3	0	100%
309	Petroleum ether	8032-32-4	III	TN	TN					TN	3	3	0	100%
326	BRIJ-35	9002-92-0	III				TN			TN	2	2	0	100%
5	Triclocarban	101-20-2	IV			TN		TN			2	2	0	100%
67	4,4'-Methylene bis-(2,6-di-tert-butylphenol)	118-82-1	IV			TN		TN			2	2	0	100%
216	1,5-Hexadiene	592-42-7	III					TN		TN	2	2	0	100%
122	Dicaprylyl carbonate	1680-31-5	IV			TN		TN			2	2	0	100%
164	Myristyl myristate	3234-85-3	IV			TN		TN			2	2	0	100%
230	Polyethylene glycol (PEG-40) hydrogenated castor oil	61788-85-0	IV			TN				TN	2	2	0	100%
251	Polyglyceryl-3 diisooctadecanoate	63705-03-3	IV			TN		TN			2	2	0	100%
324	2-Propylheptyl octanoate	868839-23-0	III			TN		TN			2	2	0	100%
201	3,3-Dimethylpentane	562-49-2	IV			TN				TN	2	2	0	100%
70	Anthracene	120-12-7		TN	TN						2	2	0	100%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	TN	FP	%
7	Hexyl cinnamic aldehyde	101-86-0	IV						TN	TN	2	2	0	100%
29	Diisobutyl ketone	108-83-8	IV						TN	TN	2	2	0	100%
38	Isopropyl myristate	110-27-0	IV						TN	TN	2	2	0	100%
47	1,5-Dibromopentane	111-24-0	III					TN	TN		2	2	0	100%
83	Xylene	1330-20-7	II						TN	TN	2	2	0	100%
157	iso-Octyl acrylate	29590-42-9	IV						TN	TN	2	2	0	100%
194	Cyclopentasiloxane	541-02-6							TN	TN	2	2	0	100%
243	1,6-Dibromohexane	629-03-8	IV					TN	TN		2	2	0	100%
285	Isopropyl bromide	75-26-3	IV						TN	TN	2	2	0	100%
204	Glycerol	56-81-5	IV	TN	TN	TN	FP	TN	TN	TN	7	6	1	86%
147	Polyethylene glycol 400	25322-68-3	IV	TN	TN	TN	FP		TN	TN	6	5	1	83%
61	Dodecane	112-40-3	III	TN	TN	TN		FP	TN	TN	6	5	1	83%
353	Methyl cyclopentane	96-37-7	III	FP	TN	TN	TN			TN	5	4	1	80%
19	1,2,6-Hexanetriol	106-69-4	IV	TN		FP			TN	TN	4	3	1	75%
345	Phenothiazine	92-84-2		FP	TN	TN					3	2	1	67%
9	Triethanolamine	102-71-6	III		TN		FP			TN	3	2	1	67%
2	Styrene	100-42-5	III					TN	FP	TN	3	2	1	67%
121	1,9-Decadiene	1647-16-1	IV	FP		TN		TN	FP	TN	5	3	2	60%
335	Tween 20	9005-64-5	III		TN	TN	FP	TN		FP	5	3	2	60%
275	1-Bromo-4-chlorobutane	6940-78-9	IV			FP		FP	TN	TN	4	2	2	50%
240	2,4-Pentanediol	625-69-4	IV	FP				FP	TN	TN	4	2	2	50%
168	p-Methyl thiobenzaldehyde	3446-89-7	IV			TN			FP		2	1	1	50%
39	Methyl amyl ketone	110-43-0	III			FP				TN	2	1	1	50%
351	1,2,4-Trimethylbenzene	95-63-6	III		FP					TN	2	1	1	50%
352	1,2,3-Trichloropropane	96-18-4	III		FP					TN	2	1	1	50%
264	Dimethyl sulfoxide	67-68-5	III	FP	FP	FP	TN			TN	5	2	3	40%
96	Cetylpyridinium bromide 0.1%	140-72-7	III		FP		FP	TN			3	1	2	33%
86	Silicic acid	1343-98-2		FP	FP	TN					3	1	2	33%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	TN	FP	%
318	Theobromine	83-67-0		FP	FP	TN					3	1	2	33%
106	2–Mercaptopyrimidine	1450-85-7	IV	FP	FP	TN					3	1	2	33%
308	Tetrabromobisphenol A	79-94-7		FP	FP	TN					3	1	2	33%
146	3-Glycidoxypropyltrimethoxysilane	2530-83-8	IV	FP	FP					TN	3	1	2	33%
174	Ethyl trimethyl acetate	3938-95-2	III		FP		FP			TN	3	1	2	33%
76	n-Butyl acetate	123-86-4	III	FP	FP		FP			TN	4	1	3	25%
175	2,2-Dimethyl-3-pentanol	3970-62-5	III			FP		FP	FP	TN	4	1	3	25%
221	N,N-Dimethylguanidine sulfate	598-65-2	III			FP		FP	FP	TN	4	1	3	25%
362	1,3-Di-iso-propylbenzene	99-62-7	IV	FP			FP	FP	FP	TN	4	1	3	25%
30	Toluene	108-88-3	III		FP	FP	FP	FP		TN	5	1	4	20%
26	Methyl isobutyl ketone	108-10-1	III	FP	FP	FP	FP			TN	5	1	4	20%
112	Sodium lauryl sulfate (3%)	151-21-3	III	FP	FP	FP	FP	FP			5	0	5	0%
101	Ethyl acetate	141-78-6	III		FP		FP	FP	FP	FP	5	0	5	0%
282	1-(9H-Carbazol-4-yloxy)-3-[[2-(2-methoxyphenoxy)ethyl]amino]propan-2-ol	72956-09-3	IV	FP	FP	FP		FP			4	0	4	0%
193	Tetraaminopyrimidine sulfate	5392-28-9	III	FP	FP		FP		FP		4	0	4	0%
32	Cyclohexanone	108-94-1	III		FP	FP		FP		FP	3	0	3	0%
349	Propyl-4-hydroxybenzoate	94-13-3	III	FP	FP	FP		FP			3	0	3	0%
55	2-(2-Ethoxyethoxy) ethanol	111-90-0	III			FP		FP	FP		3	0	3	0%
90	3-Phenoxy benzyl alcohol	13826-35-2	III			FP		FP	FP		3	0	3	0%
139	2-Ethoxyethyl methacrylate	2370-63-0	IV			FP	FP	FP	FP		3	0	3	0%
74	2,4-Pentanedione	123-54-6	III		FP				FP	FP	3	0	3	0%
20	Glycidyl methacrylate	106-91-2	III			FP	FP			FP	2	0	2	0%
71	3,4-Dimethoxy benzaldehyde	120-14-9	III		FP	FP		FP			2	0	2	0%
25	1-Nitropropane	108-03-2	IV	FP	FP						2	0	2	0%
102	Ethyl acetoacetate	141-97-9	III	FP	FP						2	0	2	0%
34	2-Methoxyethanol	109-86-4	III		FP		FP				2	0	2	0%
287	Trichloroacetic acid (3%)	76-03-9	III		FP		FP				2	0	2	0%
244	Ethylene glycol diethyl ether	629-14-1	IV					FP	FP		2	0	2	0%

Table 5A. Identification of Non-Classified Chemicals. # is the chemical number taken from Supplemental Table 2 (which provides additional information on each chemical and, for ease of use, is organized by ascending CAS Registry Number). In vivo EPA classifications from I to IV. BCOP = Bovine Corneal Opacity and Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; OI = Ocular Irritection, OS = OptiSafe; STE = Short Time Exposure; TN = True Negative; FP = False Positive; % is the number of FPs divided by the total tests performed on each chemical.

Table 5B.

Identification of Irritant Chemicals.

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	FN	TP	%
54	1-Octanol	111-87-5	II	TP	TP	TP	TP	TP	TP	TP	7	0	7	100%
49	n-Hexanol	111-27-3	II	TP	TP	TP	TP	TP	TP	TP	7	0	7	100%
44	Pyridine	110-86-1	I	TP	TP	TP	TP	TP		TP	6	0	6	100%
31	Cyclohexanol	108-93-0	I	TP	TP	TP	TP		TP	TP	6	0	6	100%
14	2-Ethyl-1-hexanol	104-76-7	II	TP	TP	TP	TP	TP		TP	6	0	6	100%
227	Ethyl-2-methyl acetoacetate	609-14-3	III	TP	TP	TP	TP		TP	TP	6	0	6	100%
263	Acetone	67-64-1	II	TP	TP	TP	TP		TP	TP	6	0	6	100%
118	Dibenzyl phosphate	1623-08-1	II	TP	TP	TP	TP	TP	TP		6	0	6	100%
354	Cyclopentanol	96-41-3	II	TP	TP	TP		TP	TP	TP	6	0	6	100%
103	Lauric Acid	143-07-7	I	TP	TP	TP		TP	TP	TP	6	0	6	100%
16	2-Methyl-1-pentanol	105-30-6	III	TP	TP	TP		TP	TP	TP	6	0	6	100%
189	Sodium benzoate	532-32-1	II	TP	TP	TP		TP	TP		5	0	5	100%
249	Benzalkonium chloride (5%)	63449-41-2/8001-54-5	I	TP	TP	TP	TP	TP			5	0	5	100%
156	Imidazole	288-32-4	I	TP	TP	TP	TP			TP	5	0	5	100%
361	p-Tert-butylphenol	98-54-4	I	TP	TP	TP		TP	TP		5	0	5	100%
301	Isobutyraldehyde	78-84-2	III	TP	TP	TP			TP	TP	5	0	5	100%
79	Sodium hydroxide (10%)	1310-73-2	I		TP	TP	TP	TP		TP	5	0	5	100%
110	Sodium lauryl sulfate (15%)	151-21-3	I		TP	TP	TP	TP		TP	5	0	5	100%
247	Benzalkonium chloride (10%)	63449-41-2/8001-54-5	I		TP	TP	TP	TP		TP	5	0	5	100%
241	Sodium oxalate	62-76-0	I		TP	TP	TP	TP	TP		5	0	5	100%
355	gamma-Butyrolactone	96-48-0	II		TP	TP		TP	TP	TP	5	0	5	100%
183	Lactic Acid	50-21-5	I	TP	TP	TP			TP	TP	5	0	5	100%
75	n-Butanal	123-72-8	III	TP	TP	TP			TP	TP	5	0	5	100%
195	Isopropyl acetoacetate	542-08-5	III	TP	TP	TP		TP	TP	TP	5	0	5	100%
211	Promethazine hydrochloride	58-33-3	I		TP		TP	TP	TP	TP	5	0	5	100%
300	Isobutanol	78-83-1	II		TP		TP	TP	TP	TP	5	0	5	100%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	FN	TP	%
15	Sodium perborate tetrahydrate	10486-00-7	I	TP	TP	TP	TP				4	0	4	100%
72	2–Benzyl–4–chlorophenol	120-32-1		TP	TP	TP		TP			4	0	4	100%
228	1–Naphthalene acetic acid Na salt	61-31-4	I	TP	TP	TP	TP				4	0	4	100%
84	Captan 90–concentrate	133-06-2	I	TP	TP	TP	TP				4	0	4	100%
138	Methyl thioglycolate	2365-48-2	II	TP	TP	TP			TP		4	0	4	100%
197	3-Chloropropionitrile	542-76-7	III	TP	TP	TP			TP		4	0	4	100%
56	3,3'–Dithiopropionic acid	1119-62-6	II	TP	TP	TP			TP		4	0	4	100%
154	Dibenzoyl–L–tartaric acid	2743-38-6	I	TP	TP	TP	TP				4	0	4	100%
359	Benzensulphonylchloride	98-09-9		TP	TP	TP				TP	4	0	4	100%
286	Trichloroacetic acid (30%)	76-03-9	I	TP	TP	TP	TP				4	0	4	100%
200	Chlorhexidine	55-56-1	I		TP	TP	TP	TP			4	0	4	100%
94	Cetylpyridinium bromide (6%)	140-72-7	I		TP	TP	TP	TP			4	0	4	100%
81	Sodium hydroxide (1%)	1310-73-2	III	TP	TP	TP	TP				4	0	4	100%
328	Triton X-100 (5%)	9002-93-1	III		TP	TP	TP			TP	4	0	4	100%
28	m–Phenylene diamine	108-45-2	I	TP	TP	TP				TP	4	0	4	100%
191	Domiphen bromide (10%)	538-71-6	I	TP	TP	TP				TP	4	0	4	100%
1	Diethylethanolamine	100-37-8	I	TP	TP	TP				TP	4	0	4	100%
35	Tetrahydrofuran	109-99-9		TP	TP	TP				TP	4	0	4	100%
53	n–Octylamine	111-86-4		TP	TP	TP				TP	4	0	4	100%
78	Ethylhexyl acid phosphate ester	12645-31-7		TP	TP	TP				TP	4	0	4	100%
312	2-Hydroxy iso–butyric acid ethyl ester	80-55-7		TP	TP	TP				TP	4	0	4	100%
343	(3–Aminopropyl)triethoxy silane	919-30-2		TP	TP	TP				TP	4	0	4	100%
205	Cetyltrimethylammonium bromide (10%)	57-09-0	I	TP	TP	TP				TP	4	0	4	100%
23	Distearyldimethylammonium chloride	107-64-2	I	TP	TP	TP				TP	4	0	4	100%
295	2–Propanol, 1–phenoxy–	770-35-4		TP	TP	TP				TP	4	0	4	100%
162	Methoxyethyl acrylate	3121-61-7	III	TP	TP	TP				TP	4	0	4	100%
303	Glycolic acid (10%)	79-14-1		TP	TP	TP				TP	4	0	4	100%
137	1-Chlorooctan-8-ol	23144-52-7		TP	TP	TP				TP	4	0	4	100%
202	Tributyltin oxide	56-35-9		TP	TP	TP				TP	4	0	4	100%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	FN	TP	%
57	Hexadecyltrimethylammonium chloride (25%)	112-02-7		TP	TP	TP				TP	4	0	4	100%
302	Methyl ethyl ketone	78-93-3	III	TP	TP		TP			TP	4	0	4	100%
93	Cetylpyridinium bromide (10%)	140-72-7	I		TP		TP	TP		TP	4	0	4	100%
278	n-Butanol	71-36-3	II				TP	TP	TP	TP	4	0	4	100%
50	n-Acetyl-DL-methionine	1115-47-5		TP	TP	TP					3	0	3	100%
347	N-(2-Methylphenyl)-iminodicarbonimidic diamide (1-(o-Tolyl)biguanide)	93-69-6		TP	TP	TP					3	0	3	100%
142	bis-(3-Aminopropyl)-tetramethyldisiloxane	2469-55-8		TP	TP	TP					3	0	3	100%
88	Triethanolamine orthovanadate	13476-99-8		TP	TP	TP					3	0	3	100%
165	alpha-Ketoglutaric acid	328-50-7		TP	TP	TP					3	0	3	100%
219	2-Hydroxy iso-butyric acid	594-61-6		TP	TP	TP					3	0	3	100%
161	Paraformaldehyde	30525-89-4		TP	TP	TP					3	0	3	100%
149	1,2-Benzisothiazol-3(2H)-one	2634-33-5		TP	TP	TP					3	0	3	100%
177	4,4'-(4,5,6,7-Tetrabromo-3H-2,1-benzoxathiol-3-ylidene)bis[2,6-dibromophenol] S,S-dioxide	4430-25-5		TP	TP	TP					3	0	3	100%
258	Benzoic acid	65-85-0		TP	TP	TP					3	0	3	100%
319	2-Hydroxy-1,4-naphthoquinone	83-72-7		TP	TP	TP					3	0	3	100%
237	4-Nitrobenzoic acid	62-23-7		TP	TP	TP					3	0	3	100%
87	n,n-Diethyl-m-toluamide	134-62-3	III			TP		TP	TP		3	0	3	100%
127	Chlorhexidine digluconate solution (20% aqueous)	18472-51-0	II			TP		TP	TP		3	0	3	100%
170	Parafluoraniline	371-40-4		TP	TP		TP				3	0	3	100%
248	Benzalkonium chloride (100%)	63449-41-2/8001-54-5	I		TP		TP			TP	3	0	3	100%
206	Cetyltrimethylammonium bromide	57-09-0		TP	TP					TP	3	0	3	100%
250	Benzalkonium chloride (1%)	63449-41-2/8001-54-5	I		TP		TP	TP			3	0	3	100%
329	Triton X-100 (10%)	9002-93-1	II		TP	TP	TP			TP	3	0	3	100%
213	Potassium cyanate	590-28-3		TP	TP	TP					3	0	3	100%
255	Acetic Acid (10%)	64-19-7	I	TP			TP		TP		3	0	3	100%
108	Tetraoctylammonium bromide	14866-33-2	I			TP			TP		2	0	2	100%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	FN	TP	%
323	1-Naphthalene acetic acid	86-87-3				TP	TP				2	0	2	100%
51	Butyl cellosolve	111-76-2	II		TP					TP	2	0	2	100%
21	Allyl alcohol	107-18-6	III		TP				TP		2	0	2	100%
73	Benzethonium chloride	121-54-0	I		TP					TP	2	0	2	100%
160	Sodium deoxycholate (10%)	302-95-4	II		TP			TP			2	0	2	100%
220	2,2-Dimethylbutanoic acid	595-37-9	I		TP		TP				2	0	2	100%
3	Benzyl alcohol	100-51-6	I	TP						TP	2	0	2	100%
109	Sodium lauryl sulfate	151-21-3	III				TP			TP	2	0	2	100%
208	Butanedioic acid, sulfo-, 1, 4-bis(2-ethylhexyl) ester, sodium salt	577-11-7	I						TP	TP	2	0	2	100%
304	Glycolic acid	79-14-1	I						TP	TP	2	0	2	100%
125	(Ethylenediamine propyl)trimethoxysilane	1760-24-3		TP		TP					2	0	2	100%
256	Ammonium nitrate	6484-52-2	III	TP	TP	TP	TP	TP	TP	FN	7	1	6	86%
262	Isopropanol	67-63-0	III	TP	TP	TP	TP	TP	TP	FN	7	1	6	86%
305	Methyl acetate	79-20-9	II	TP	TP	TP	TP	TP	TP	FN	7	1	6	86%
115	1-Propoxy-2-propanol (Propasol solvent P)	1569-01-3	II	TP	TP	TP		TP	TP	FN	6	1	5	83%
173	Sodium chloroacetate	3926-62-3	III	TP	FN	TP		TP	TP	TP	6	1	5	83%
317	1,5–Naphthalenediol	83-56-7	II	TP	TP	TP		FN	TP		5	1	4	80%
37	2,5-Dimethyl-2,5-hexanediol	110-03-2	I		TP	TP	TP		TP	FN	5	1	4	80%
253	Ethanol	64-17-5	I		TP	TP	TP	TP		FN	5	1	4	80%
234	4–Carboxybenzaldehyde	619-66-9	II	TP	TP	TP	FN				4	1	3	75%
314	Hydroxyethyl acrylate	818-61-1		TP	TP	TP				FN	4	1	3	75%
196	Sodium salicylate	54-21-7	I	TP	TP	TP				FN	4	1	3	75%
158	1-(1-Methyl-2-propoxyethoxy)propan-2-ol	29911-27-1	III	TP	TP	TP				FN	4	1	3	75%
179	2,6-Dichlorobenzoyl chloride	4659-45-4	II	TP	FN	TP	TP		TP	FN	6	2	4	67%
123	2-Amino-3-pyridinol	16867-03-1	III	TP	FN	TP					3	1	2	67%
363	m–Dinitrobenzene	99-65-0		TP	FN	TP					3	1	2	67%
356	Ethyl 2,6-dichloro-5-fluoro-beta-oxo-3-pyridine propionate	96568-04-6		FN	TP	TP					3	1	2	67%
126	Tetraethylene glycol diacrylate	17831-71-9	I			TP	FN	TP			3	1	2	67%

#	Chemical Name	CASRN	in vivo EPA	BCOP (LLBO)	BCOP (OP-KIT)	EPI	ICE	OI	OS	STE	Number of tests	FN	TP	%
17	Methyl cyanoacetate	105-34-0	II	TP	TP	FN	FN	TP	TP	FN	7	3	4	57%
307	Camphene	79-92-5	III	FN	FN	TP		TP	TP	FN	6	3	3	50%
91	4–(1,1,3,3–Tetramethylbutyl)phenol	140-66-9		FN	FN	TP		TP			4	2	2	50%
274	Quinacrine	69-05-6	I	FN	FN	TP	TP				4	2	2	50%
98	2–Pseudoionone	141-10-6		TP	FN	FN				TP	4	2	2	50%
291	N-Lauroylsarcosine, sodium salt (10%)	7631-98-3/137-16-6	III		TP			FN			2	1	1	50%
62	Myristyl alcohol	112-72-1	III						TP	FN	2	1	1	50%
77	Maneb	12427-38-2	III				FN		TP		2	1	1	50%
277	L-Aspartic acid	70-47-3		FN	FN		TP				3	2	1	33%
293	(2R,3R)-3-((R)-1-(tert-butyldimethylsilyloxy)ethyl)-4-oxoazetidin-2-yl acetate	76855-69-1		FN	FN	TP					3	2	1	33%

Table 5B. Identification of nonclassified chemicals. # is the chemical number taken from Supplemental Table 2 (which provides additional information on each chemical and, for ease of use, is organized by ascending CAS Registry Number). In vivo EPA classifications from I to IV. BCOP = Bovine Corneal Opacity and Permeability; LLBO = Laser Light-Based Opacitometer; ICE = Isolated Chicken Eye; OI = Ocular Irritection; OS = OptiSafe; STE = Short Time Exposure; TP = True Positive; FN = False Negative; % is the number of FNs divided by the total tests performed on each chemical.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 6.

Tiered Testing

Table 6A. Tiered Testing (EpiOcular and ICE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
EpiOcular	9	34	43	2.9%	44.4%	88.4%	76.3%
ICE	9	34	43	5.9%	66.7%	81.4%	63.7%
Tiered Testing (Theoretical)	9	34	43	8.7%	29.6%	86.9%	80.8%
Tiered Testing (Actual)	9	34	43	5.9%	33.3%	88.4%	80.4%

Table 6D. Tiered Testing (OI and ICE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
ICE	7	21	28	4.8%	100.0%	71.4%	47.6%
OI	7	21	28	0.0%	42.9%	89.3%	78.6%
Tiered Testing (Theoretical)	7	21	28	4.8%	42.9%	85.7%	76.2%
Tiered Testing (Actual)	7	21	28	4.8%	42.9%	85.7%	76.2%

Table 6G. Tiered Testing (EpiOcular and STE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
EpiOcular	29	51	80	3.9%	34.5%	85.0%	80.8%
STE	29	51	80	23.5%	10.3%	81.3%	83.1%
Tiered Testing (Theoretical)	29	51	80	26.5%	3.6%	81.8%	85.0%
Tiered Testing (Actual)	29	51	80	25.5%	6.9%	81.3%	83.8%

Table 6J. Tiered Testing (OS and BCOP (OPKIT))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (OPKIT)	9	34	43	8.8%	33.3%	86.0%	78.9%
OS	9	34	43	0.0%	33.3%	93.0%	83.3%
Tiered Testing (Theoretical)	9	34	43	8.8%	11.1%	90.7%	90.1%
Tiered Testing (Actual)	9	34	43	8.8%	33.3%	86.0%	78.9%

Table 6M. Tiered Testing (ICE and BCOP (OPKIT))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (OPKIT)	17	44	61	6.8%	64.7%	77.0%	64.2%
ICE	17	44	61	4.5%	82.4%	73.8%	56.6%
Tiered Testing (Theoretical)	17	44	61	11.0%	53.3%	77.2%	67.8%

Table 6B. Tiered Testing (EpiOcular and OI)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
EpiOcular	28	30	58	3.3%	35.7%	81.0%	80.5%
OI	28	30	58	6.7%	39.3%	77.6%	77.0%
Tiered Testing (Theoretical)	28	30	58	9.8%	14.0%	88.2%	88.1%
Tiered Testing (Actual)	28	30	58	10.0%	35.7%	77.6%	77.1%

Table 6E. Tiered Testing (OS and ICE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
ICE	4	20	24	10.0%	75.0%	79.2%	57.5%
OS	4	20	24	0.0%	50.0%	91.7%	75.0%
Tiered Testing (Theoretical)	4	20	24	10.0%	37.5%	85.4%	76.3%
Tiered Testing (Actual)	4	20	24	10.0%	50.0%	83.3%	70.0%

Table 6H. Tiered Testing (ICE and STE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
ICE	13	28	41	3.6%	69.2%	75.6%	63.6%
STE	13	28	41	25.0%	15.4%	78.0%	79.8%
Tiered Testing (Theoretical)	13	28	41	27.7%	10.7%	77.7%	80.8%
Tiered Testing (Actual)	13	28	41	25.0%	15.4%	78.0%	79.8%

Table 6K. Tiered Testing (BCOP (OPKIT) and STE)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (OPKIT)	24	60	84	6.7%	50.0%	81.0%	71.7%
STE	23	61	84	19.7%	17.4%	81.0%	81.5%
Tiered Testing (Theoretical)	24	60	84	25.1%	8.7%	79.6%	83.1%
Tiered Testing (Actual)	24	60	84	21.7%	12.5%	79.8%	82.9%

Table 6N. Tiered Testing (BCOP (LLBO) and BCOP (OPKIT))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (LLBO)	29	83	112	7.2%	58.6%	79.5%	67.1%
BCOP (OPKIT)	29	83	112	12.0%	51.7%	77.7%	68.1%
Tiered Testing (Theoretical)	29	83	112	18.4%	30.3%	78.5%	75.7%

Table 6C. Tiered Testing (OS and EpiOcular)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
EpiOcular	21	35	56	2.9%	33.3%	85.7%	81.9%
OS	21	35	56	0.0%	33.3%	87.5%	83.3%
Tiered Testing (Theoretical)	21	35	56	2.9%	11.1%	94.0%	93.0%
Tiered Testing (Actual)	21	35	56	2.9%	23.8%	89.3%	86.7%

Table 6F. Tiered Testing (OS and OI)

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
OI	23	23	46	8.7%	47.8%	71.7%	71.7%
OS	23	23	46	0.0%	43.5%	78.3%	78.3%
Tiered Testing (Theoretical)	23	23	46	8.7%	20.8%	85.3%	85.3%
Tiered Testing (Actual)	23	23	46	8.7%	34.8%	78.3%	78.3%

Table 6I. Tiered Testing (OI and BCOP (OPKIT))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (OPKIT)	11	33	44	9.1%	45.5%	81.8%	72.7%
OI	11	33	44	6.1%	45.5%	84.1%	74.2%
Tiered Testing (Theoretical)	11	33	44	14.6%	20.7%	83.9%	82.4%
Tiered Testing (Actual)	11	33	44	15.2%	36.4%	79.5%	74.2%

Table 6L. Tiered Testing (EpiOcular and BCOP (OPKIT))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (OPKIT)	23	88	111	10.2%	47.8%	82.0%	71.0%
EpiOcular	23	88	111	2.3%	30.4%	91.9%	83.6%
Tiered Testing (Theoretical)	23	88	111	12.2%	14.5%	87.3%	86.6%
Tiered Testing (Actual)	23	88	111	11.4%	30.4%	84.7%	79.1%

Table 6O. Tiered Testing (EpiOcular and BCOP (LLBO))

Tiered Test	In vivo NC	Cat 1 or 2	n	FN Rate	FP Rate	Acc.	Bal. Acc.
BCOP (LLBO)	25	79	104	6.3%	48.0%	83.7%	72.8%
EpiOcular	25	79	104	2.5%	24.0%	92.3%	86.7%
Tiered Testing (Theoretical)	25	79	104	8.7%	11.5%	90.6%	89.9%

Table 6P. Tiered Testing (OS and BCOP (LLBO))

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
Tiered Testing (Actual)	17	44	61	11.4%	58.8%	75.4%	64.9%
BCOP (LLBO)	13	28	41	3.6%	30.8%	87.8%	82.8%
OS	13	28	41	0.0%	23.1%	92.7%	88.5%
Tiered Testing (Theoretical)	13	28	41	3.6%	7.1%	95.3%	94.6%
Tiered Testing (Actual)	13	28	41	3.6%	23.1%	90.2%	86.7%

Table 6S. Tiered Testing (BCOP (LLBO) and STE)

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
BCOP (LLBO)	22	47	69	2.1%	40.9%	85.5%	78.5%
STE	21	48	69	20.8%	0.0%	85.5%	89.6%
Tiered Testing (Theoretical)	22	47	69	22.5%	0.0%	84.7%	88.8%
Tiered Testing (Actual)	22	47	69	19.1%	0.0%	85.5%	90.4%

Table 6Q. Tiered Testing (BCOP (LLBO) and ICE)

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
Tiered Testing (Actual)	29	83	112	13.3%	51.7%	76.8%	67.5%
BCOP (LLBO)	9	28	37	7.1%	66.7%	78.4%	63.1%
ICE	9	28	37	7.1%	66.7%	78.4%	63.1%
Tiered Testing (Theoretical)	9	28	37	13.7%	44.5%	78.8%	70.9%
Tiered Testing (Actual)	9	28	37	14.3%	44.4%	78.4%	70.6%

Table 6T. Tiered Testing (OI and STE)

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
OI	19	21	40	0.0%	42.1%	80.0%	78.9%
STE	19	21	40	33.3%	10.5%	77.5%	78.1%
Tiered Testing (Theoretical)	19	21	40	33.3%	4.4%	80.4%	81.1%
Tiered Testing (Actual)	19	21	40	33.3%	5.3%	80.0%	80.7%

Table 6R. Tiered Testing (BCOP (LLBO) and OI)

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
Tiered Testing (Actual)	25	79	104	8.9%	20.0%	88.5%	85.6%
BCOP (LLBO)	11	21	32	9.5%	45.5%	78.1%	72.5%
OI	11	21	32	4.8%	45.5%	81.3%	74.9%
Tiered Testing (Theoretical)	11.03	21	32	13.8%	20.7%	83.8%	82.7%
Tiered Testing (Actual)	11	21	32	14.3%	36.4%	78.1%	74.7%

Table 6U. Tiered Testing (OS and STE)

Tiered Test	In vivo		n	FN Rate	FP Rate	Acc.	Bal. Acc.
	NC	Cat 1 or 2					
OS	27	29	56	0.0%	25.9%	87.5%	87.0%
STE	27	29	56	31.0%	7.4%	80.4%	80.8%
Tiered Testing (Theoretical)	27	29	56	31.0%	1.9%	83.0%	83.5%
Tiered Testing (Actual)	27	29	56	31.0%	7.4%	80.4%	80.8%

Table 6. Example of sequential testing strategy with test 1 first and then test 2, using chemicals in common between both tests. Materials are first assayed using test 1. Since these results are available, the accuracy values are the actual values, which are also the same as the theoretical accuracy values for the chemical set. In this sequential strategy, only positive results will be re-tested using test 2 (chemicals in the data set that were identified as in vitro positives will be retested using test 2). Test 1 then test 2 (theoretical) assumes the accuracy for the second test is as described in Table 2 (accuracy comparison table); in other words, accuracy is unaffected by the outcome of the first test. Test 1 then test 2 (actual) shows the real accuracy values. Since the results for all of these materials are known for both tests, the actual outcome of the first test is followed by the sequential testing of positives by the second test (refer to Supplemental Table 2).