13.0 GLOSSARY[1]

Accuracy[2]: (a) The closeness of agreement between a test method result and an accepted reference value. (b) The proportion of correct outcomes of a test method. It is a measure of test method performance and one aspect of "relevance." The term is often used interchangeably with "concordance" (see also "two-by-two" table). Accuracy is highly dependent on the prevalence of positives in the population being examined.

Assay[2]: The experimental system used. Often used interchangeably with "test" and "test method."

Benchmark substance: A substance used as a standard for comparison to a test substance. A benchmark substance should have the following properties:

- a consistent and reliable source(s)
- structural and functional similarity to the class of substances being tested
- known physical/chemical characteristics
- supporting data on known effects
- known potency in the range of the desired response

Benchmark control: A sample containing all components of a test system and treated with a known substance (i.e., the benchmark substance) to induce a known response. The sample is processed with test substance-treated and other control samples to compare the response produced by the test substance to the benchmark substance to allow for an assessment of the sensitivity of the test method to assess a specific chemical class or product class.

Blepharitis: Inflammation of the eyelids.

Bulbar conjunctiva: The portion of the conjunctiva that covers the outer surface of the eye.

Chemosis: A form of eye irritation in which the membranes that line the eyelids and surface of the eye ("conjunctiva") become swollen.

Classification system: An arrangement of quantified results or data into groups or categories according to previously established criteria.

Coded substances: Substances labeled by code rather than name so that they can be tested and evaluated without knowledge of their identity or anticipation of test results. Coded substances are used to avoid intentional or unintentional bias when evaluating laboratory or test method performance.

[1] The definitions in this Glossary are restricted to their uses with respect to the Draize rabbit eye test method and the IRE test method.

Coefficient of variation: A statistical representation of the precision of a test. It is expressed as a percentage and is calculated as follows:

$$\left(\frac{standard\ deviation}{mean}\right) \times 100\%$$

Concordance[2]: The proportion of all substances tested that are correctly classified as positive or negative. It is a measure of test method performance and one aspect of "relevance". The term is often used interchangeably with "accuracy" (see also "two-by-two" table). Concordance is highly dependent on the prevalence of positives in the population being examined.

Conjunctiva: The mucous membrane that lines the inner surfaces of the eyelids and folds back to cover the front surface of the eyeball, except for the central clear portion of the outer eye (the cornea). The conjunctiva is composed of three sections: palpebral conjunctiva, bulbar conjunctiva, and fornix.

Conjunctival sac: The space located between the eyelid and the conjunctiva-covered eyeball. Substances are instilled into the sac to conduct an *in vivo* eye test.
Cornea: The transparent part of the coat of the eyeball that covers the iris and pupil and admits light to the interior.

Corneal opacity: Measurement of the extent of opaqueness of the cornea following exposure to a test substance. Increased corneal opacity is indicative of damage to the cornea. Opacity can be evaluated subjectively as done in the Draize rabbit eye test, or objectively with an instrument such as an "opacitometer."

Corneal permeability: Quantitative measurement of damage to the corneal epithelium by a determination of the amount of sodium fluorescein dye that passes through all corneal cell layers.

Corneal swelling: An objective measurement in the IRE test of the extent of distention of the cornea following exposure to a test substance. It is expressed as a percentage and is calculated from corneal thickness measurements that are recorded at regular intervals during the IRE test. Increased corneal swelling is indicative of damage to the corneal epithelium.

Corneal thickness: The depth of the cornea measured using an ultrasonic pachymeter or a depth-measuring attachment on a slit-lamp.

Corrosion: Destruction of tissue at the site of contact with a substance.

Corrosive: A substance that causes irreversible tissue damage at the site of contact.

Endpoint[2]: The biological process, response, or effect assessed by a test method.

Enucleate: To remove without cutting into.

False negative[2]: A substance incorrectly identified as negative by a test method.

False negative rate[2]: The proportion of all positive substances falsely identified by a test method as negative (see "two-by-two" table). It is one indicator of test method accuracy.

False positive[2]: A substance incorrectly identified as positive by a test method.

False positive rate[2]: The proportion of all negative substances that are falsely identified by a test method as positive (see "two-by-two" table). It is one indicator of test method accuracy.

Fibrous tunic: The outer of the three membranes of the eye, comprising the cornea and the sclera; also called *tunica fibrosa oculi*.

Fluorescein penetration/retention: A subjective measurement in the IRE test of the extent of fluorescein sodium that is retained by epithelial cells in the cornea following exposure to a test substance. Increased fluorescein retention is indicative of damage to the corneal epithelium.

Globally Harmonized System (GHS): A classification system presented by the United Nations that provides (a) a harmonized criteria for classifying substances and mixtures according to their health, environmental and physical hazards, and (b) harmonized hazard communication elements, including requirements for labeling and safety data sheets.

Good Laboratory Practices (GLP)[2]: Regulations promulgated by the U.S. Food and Drug Administration and the U.S. Environmental Protection Agency, and principles and procedures adopted by the Organization for Economic Cooperation and Development and Japanese authorities that describe record keeping and quality assurance procedures for laboratory records that will be the basis for data submissions to national regulatory agencies.

Hazard[2]: The potential for an adverse health or ecological effect. A hazard potential results only if an exposure occurs that leads to the possibility of an adverse effect being manifested.

Interlaboratory reproducibility[2]: A measure of whether different qualified laboratories using the same protocol and test substances can produce qualitatively and quantitatively similar results. Interlaboratory reproducibility is determined during the prevalidation and validation processes and indicates the extent to which a test method can be transferred successfully among laboratories.

Intralaboratory repeatability[2]: The closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time period.

Intralaboratory reproducibility[2]: The first stage of validation; a determination of whether qualified people within the same laboratory can successfully replicate results using a specific test protocol at different times.

In vitro: In glass. Refers to assays that are carried out in an artificial system (e.g., in a test tube or petri dish) and typically use single-cell organisms, cultured cells, cell-free extracts, or purified cellular components.

In vivo: In the living organism. Refers to assays performed in multicellular organisms.

Iris: The contractile diaphragm perforated by the pupil and forming the colored portion of the eye.

Negative control: An untreated sample containing all components of a test system, except the test substance solvent, which is replaced with a known nonreactive material, such as water. This sample is processed with test substance-treated samples and other control samples to determine whether the solvent interacts with the test system.

Negative predictivity[2]: The proportion of correct negative responses among substances testing negative by a test method (see "two-by-two" table). It is one indicator of test method accuracy. Negative predictivity is a function of the sensitivity of the test method and the prevalence of negatives among the substances tested.

Neuroectodermal tunic: The innermost of three membranes of the eye, comprising the retina.

Nictating (nictitating) membrane: The membrane that moves horizontally across the eye in some animal species (e.g., rabbit, cat) to provide additional protection in particular circumstances. It may be referred to as the "third eyelid."

Nonirritant: (a) A substance the produces no changes in the eye following application to the anterior surface of the eye. (b) Substances that are not classified as GHS Category 1, 2A, or 2B; or EU R41 or R36 ocular irritants.

Nonsevere irritant: (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye; the tissue damage is reversible within 21 days of application and the observed adverse effects in the eye are less severe than observed for a severe irritant. (b) Substances that are classified as GHS Category 2A or 2B; EPA Category II, III, or IV; or EU R36 ocular irritants.

Ocular: Of or relating to the eye.

Ocular corrosive: A substance that causes irreversible tissue damage in the eye following application to the anterior surface of the eye.

Ocular irritant: A substance that produces a reversible change in the eye following application to the anterior surface of the eye.

Opacitometer: An instrument used to measure "corneal opacity" by quantitatively evaluating light transmission through the cornea. The instrument has two compartments, each with its own light source and photocell. One compartment is used for the treated cornea, while the other is used to calibrate and zero the instrument. The difference between photocell signals in the two compartments is measured electronically as a change in voltage, and is displayed digitally, generating numerical opacity values with arbitrary units.

Palpebral conjunctiva: The part of the conjunctiva that covers the inner surface of the eyelids.

Pannus: A specific type of corneal inflammation that begins within the conjunctiva, and with time spreads to the cornea. Also referred to as "chronic superficial keratitis."

Performance[2]: The accuracy and reliability characteristics of a test method (see "accuracy", "reliability").

pH: A measure of the acidity or alkalinity of a solution. pH 7.0 is neutral; higher pHs are alkaline, lower pHs are acidic.

Positive control: A sample containing all components of a test system and treated with a substance known to induce a positive response, which is processed with the test substance-treated and other control samples to demonstrate the sensitivity of each experiment and to allow for an assessment of variability in the conduct of the assay over time.

Positive predictivity[2]: The proportion of correct positive responses among substances testing positive by a test method (see "two-by-two" table). It is one indicator of test method accuracy. Positive predictivity is a function of the sensitivity of the test method and the prevalence of positives among the substances tested.

Prevalence[2]: The proportion of positives in the population of substances tested (see "two-by-two" table).

Protocol[2]: The precise, step-by-step description of a test method, including a listing of all necessary reagents, criteria and procedures for evaluation of the test data.

Quality assurance[2]: A management process by which adherence to laboratory testing standards, requirements, and record keeping procedures is assessed independently by individuals other than those performing the testing.

Reduction alternative[2]: A new or modified test method that reduces the number of animals required.

Reference test method[2]: The accepted *in vivo* test method used for regulatory purposes to evaluate the potential of a test substance to be hazardous to the species of interest.

Refinement alternative[2]: A new or modified test method that refines procedures to lessen or eliminate pain or distress in animals or enhances animal well-being.

Relevance[2]: The extent to which a test method correctly predicts or measures the biological effect of interest in humans or another species of interest. Relevance incorporates consideration of the "accuracy" or "concordance" of a test method.

Reliability[2]: A measure of the degree to which a test method can be performed reproducibly within and among laboratories over time. It is assessed by calculating intra- and inter-laboratory reproducibility and intralaboratory repeatability.

Replacement alternative[2]: A new or modified test method that replaces animals with nonanimal systems or one animal species with a phylogenetically lower one (e.g., a mammal with an invertebrate).

Reproducibility[2]: The consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test substances (see intra- and inter-laboratory reproducibility).

Sclera: The tough, fibrous tissue that extends from the cornea to the optic nerve at the back of the eye.

Sensitivity[2]: The proportion of all positive substances that are classified correctly as positive in a test method. It is a measure of test method accuracy (see "two-by-two" table).

Secondary bacterial keratitis: Inflammation of the cornea that occurs secondary to another insult that compromised the integrity of the eye.

Slit-lamp microscope: An instrument used to directly examine the eye under the magnification of a binocular microscope by creating a stereoscopic, erect image. In the IRE test method, this instrument is used to view the anterior structures of the rabbit eye as well as to objectively measure corneal thickness with a depth-measuring device attachment.

Severe irritant: (a) A substance that causes tissue damage in the eye following application to the anterior surface of the eye that is not reversible within 21 days of application or causes serious physical decay of vision. (b) Substances that are classified as GHS Category 1, EPA Category I, or EU R41 ocular irritants.

Solvent control: An untreated sample containing all components of a test system, including the solvent that is processed with the test substance-treated and other control samples to establish the baseline response for the samples treated with the test substance dissolved in the

same solvent. When tested with a concurrent negative control, this sample also demonstrates whether the solvent interacts with the test system.

Specificity[2]: The proportion of all negative substances that are classified correctly as negative in a test method. It is a measure of test method accuracy (see "two-by-two" table).

Superfusion apparatus: Water-jacketed, temperature-controlled, custom-built apparatus usually made from Perspex® plastic that houses isolated rabbit eyes in removable holders placed in isolated chambers to provide short-term maintenance of metabolic and physiological activity. The chambers have darkened walls to permit slit-lamp examination, and drip tubes positioned over the eyes to provide a continuous saline flow to maintain the eyes during an experiment.

Test[2]: The experimental system used; used interchangeably with "test method" and "assay."

Test method[2]: A process or procedure used to obtain information on the characteristics of a substance or agent. Toxicological test methods generate information regarding the ability of a substance or agent to produce a specified biological effect under specified conditions. Used interchangeably with "test" and "assay." See also "validated test method" and "reference test."

Test method component: Structural, functional, and procedural elements of a test method that are used to develop the test method protocol. These components include unique characteristics of the test method, critical procedural details, and quality control measures.

Tiered testing: A testing strategy where all existing information on a test substance is reviewed, in a specified order, prior to *in vivo* testing. If the irritancy potential of a test substance can be assigned, based on the existing information, no additional testing is required. If the irritancy potential of a test substance cannot be assigned, based on the existing information, a step-wise animal testing procedure is performed until an unequivocal classification can be made.

Toxic keratoconjunctivitis: Inflammation of the cornea and conjunctiva due to contact with an exogenous agent. Used interchangeably with "contact keratoconjunctivitis, irritative keratoconjunctivitis, and chemical keratoconjunctivitis."

Transferability[2]: The ability of a test method or procedure to be accurately and reliably performed in different, competent laboratories.

Two-by-two table[2]: The two-by-two table can be used for calculating accuracy (concordance) ([a+d]/[a+b+c+d]), negative predictivity (d/[c+d]), positive predictivity (a/[a+b]), prevalence ([a+c]/[a+b+c+d]), sensitivity (a/[a+c]), specificity (d/[b+d]), false positive rate (b/[b+d]), and false negative rate (c/[a+c]).

		New Test Outcome		
		Positive	**Negative**	**Total**
Reference Test Outcome	**Positive**	a	c	a + c
	Negative	b	d	b + d
	Total	a + b	c + d	a + b + c + d

Uvea tract: The middle of three membranes of the eye, comprising the iris, ciliary body, and choroid. Also referred to as the "vascular tunic."

Validated test method[2]: An accepted test method for which validation studies have been completed to determine the relevance and reliability of this method for a specific proposed use.

Validation[2]: The process by which the reliability and relevance of a procedure are established for a specific purpose.

Vascular tunic: The middle of three membranes of the eye, comprising the iris, ciliary body, and choroid. Also referred to as the "uvea."

Weight of evidence (process): The strengths and weaknesses of a collection of information are used as the basis for a conclusion that may not be evident from the individual data.