Toxicology 327 (2015) 32–46

Contents lists available at ScienceDirect

Toxicology

journal homepage: www.elsevier.com/locate/toxicol

Review

An overview of current techniques for ocular toxicity testing

Samantha L. Wilson [a,*], Mark Ahearne [b,c], Andrew Hopkinson [a]

[a] Academic Ophthalmology, Division of Clinical Neuroscience, University of Nottingham, Queen's Medical Centre Campus NG7 2UH, UK
[b] Trinity Centre for Bioengineering, Trinity Biomedical Sciences Institute, Trinity College Dublin, Ireland, UK
[c] Department of Mechanical and Manufacturing Engineering, School of Engineering, Trinity College Dublin, Ireland, UK

ARTICLE INFO

Article history:
Received 25 September 2014
Received in revised form 5 November 2014
Accepted 6 November 2014
Available online 11 November 2014

Keywords:
Corneal equivalents
Cytotoxicity
Draize
In vitro alternatives
Organotypic models
Regulation
Validation

ABSTRACT

Given the hazardous nature of many materials and substances, ocular toxicity testing is required to evaluate the dangers associated with these substances after their exposure to the eye. Historically, animal tests such as the Draize test were exclusively used to determine the level of ocular toxicity by applying a test substance to a live rabbit's eye and evaluating the biological response. In recent years, legislation in many developed countries has been introduced to try to reduce animal testing and promote alternative techniques. These techniques include *ex vivo* tests on deceased animal tissue, computational models that use algorithms to apply existing data to new chemicals and *in vitro* assays based on two dimensional (2D) and three dimensional (3D) cell culture models. Here we provide a comprehensive overview of the latest advances in ocular toxicity testing techniques, and discuss the regulatory framework used to evaluate their suitability.

Contents

Abbreviations: BCOP, Bovine Cornea Opacity Permeability (test/assay); CAMVA, Chorioallantoic membrane vascular assay; CEET, Chicken enucleated eye test; CK, Cytokeratin; CM, Cytosensor Microphysiometer (test); CPSC, Consumer Product Safety Commission; CTPA, Cosmetic, Toiletry and Perfumery Association; DMSO, Dimethyl sulfoxide; DNA, Deoxyribonucleic acid; EC, European Commission; ECVAM, European Centre for the Validation of Alternative Methods; EET, Enucleated eye test; EIT, Eye irritation test; EURL-ECVAM, European Union Reference Laboratory for Alternatives to Animal Tests (formally ECVAM); FDA, Food and Drug Administration; FL, Fluorescein leakage; GHS, Globally Harmonized System (of classification); HCE, Human corneal epithelium; HET, Hen's egg test; HET-CAM, Hühner-embryonen test on chorioallantoic membrane; HO, Home Office (British); ICCVAM, Interagency Coordinating Committee on the Validation of Alternative Methods; ICE, Isolated chicken eye (test); IRAG, Interagency Regulatory Alternatives Group; IRE, Isolated rabbit eye (test); IVIS, *In vitro* irritancy score; JaCVAM, Japanese Centre for the Validation of Alternative Methods; LDH, Lactate dehydrogenase (leakage); LVET, Low volume eye irritation test; MAS, Maximum average score; MDCK, Madin–Dardy canine kidney (cells); MTT, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; NICEATM, NTP Interagency Center for the Evaluation of Alternative Toxicological Methods; NIEHS, National Committee of Environmental Health Sciences; NIOSHA, National Institute for Occupational Safety and Health Administartion; NRC, National Research Council; NTP, National Toxicology Programme; NZW, New Zealand White (rabbits); OECD, Organization for the Economic Co-operation and Countries Development; PLLBOA, Prototype laser light based opacitometer; QSAR, Quantitative structure–activity relationship; RCE, Rabbit corneal epithelium; REET, Rabbit enucleated eye test; SIRC, Statens Seruminstitut rabbit corneal (cells); SMI, Slug mucosal irritation (assay/test); STE, Short time exposure (test); TG, Test guidance; UN, United Nations; VMP, Validation Management Group.

* Corresponding author. Tel.: +44 7976 063 762.
E-mail address: Samantha.wilson@nottigham.ac.uk (S.L. Wilson).

http://dx.doi.org/10.1016/j.tox.2014.11.003

1. Introduction

The location, physiological structure and sensitivity of the ocular surface predispose it to exposure from a variety of potentially hazardous environmental conditions and substances on a daily basis. Many different materials and chemicals can result in damage to the cornea that may vary from irritation and inflammation causing mild discomfort to tissue corrosion resulting in irreversible blindness. These include household, industrial, agricultural and military products, cosmetics, toiletries and may even include certain ocular drugs and pharmaceuticals if incorrectly administered (Wilhelmus, 2001). While exposure to such substances may be incidental, accidental or intentional (Vinardell and Mitjans, 2008), most ocular incidents involve accidental exposure either in the workplace or at home *via* splashing with concentrated solutions, such as bleach or detergents, followed by rapid washing with water or removal *via* lacrimation (Shaw et al., 1991). To reduce the risk of exposure to dangerous substances all manufactured consumer products and their ingredients must be tested and their eye irritation potential assessed so that the public can be assured of their safety, or warned of the associated dangers. Eye toxicity tests are therefore required to ensure that the risks associated with products meet suitable safety criteria and are clearly labeled.

Historically, as toxicology testing has become more common, its reliance upon animal use has increased. This has primarily been due to the absence of more sophisticated assessment techniques and the low status of animals in society (Stephens and Mak, 2013). Ethical reconsideration of animal use for toxicology studies was driven by the emergence of the animal rights movement in the 1950s (Stephens and Mak, 2013) and its criticism of animal experimentation, in particular the use of Draize testing for cosmetics testing. In 1959, Russell and Burch performed a study based upon the philosophical concept of humanity, in which they observed that some biological experiments could be classed as "inhumane" based upon the levels of pain, distress and lasting harm experienced by the test animals (Russell et al., 1959). Their research provided the systematic basis of the 3R's: Replace, Reduce and Refine the use of sentient beings in experimental biology. This led to a general expansion of funding sources for *ex vivo* and *in vitro* alternative methods, to reduce the dependency on live animal testing, whilst also creating a political climate whereby alternative procedures were incorporated into federal and government legislation (Stephens and Mak, 2013). In this review, we will provide an overview of established and newly developed ocular toxicity tests and discuss their advantages and potential limitations.

2. Draize testing

Live animals have been used to assess and evaluate potentially harmful products to the eyes since the 18th century (Wilhelmus, 2001). The international standard assay for acute ocular toxicity is the rabbit *in vivo* Draize eye test (Draize et al., 1944) which was developed in the 1940s by the Food and Drugs Administration (FDA) in response to new laws implemented following permanent eye injuries occurring due to cosmetics use in the 1930s (Calabrese, 1987). Draize testing is a government endorsed protocol accepted by the Organization for Economic Co-operation and Countries Development (OECD, test guidance [TG] 405) (Huhtala et al., 2008; OECD, 2012b). New Zealand white (NZW) rabbits are most commonly used as they have large eyes with a well described anatomy and physiology, are easy to handle, readily available and are relatively inexpensive (Wilhelmus, 2001). The procedure involves the application of 0.1 ml (or 0.1 g solid) test substance onto the cornea and conjunctival sac of one eye of a conscious rabbit for up to 72 h while the other eye serves as an untreated control (Draize et al., 1944). The original Draize protocol used at least six rabbits per test, but this was reduced to three animals or a single animal when serious ocular damage is expected, with those with severe lesions being humanely euthanized. The latest Draize test guidelines include the application and delivery of analgesics and anesthetics (OECD, 2012b) to reduce animal pain and suffering. Rabbits are observed at selected intervals for up to 21 days for signs of irritation including redness, swelling, cloudiness, edema, hemorrhage, discharge and blindness (Huhtala et al., 2008). In cases where severe eye irritation or pain is observed, it is recommended that the animals are euthanized or removed from the study prior to the 21 day time point (OECD, 2012b). The observed degree of irritancy allows for chemicals to be classified, based on subjective scoring of the effect on the cornea, conjunctiva and iris, ranging from non-irritating to severely irritating. In fact, Draize testing is the only test formally accepted and validated to assess the full range of irritation severity. Both irreversible and reversible ocular effects can be identified using this test (Barile, 2010). Eye irritation was traditionally summarized as a "maximum average score" (MAS) which is an average value primarily focused on corneal injury, for individual animals at the time of scoring (Huhtala et al., 2008). However, many countries had their own scoring systems, which although similar in their approach, led to multiple classifications, labels, and data sheets for the same chemical, dependent upon which country the chemical was been marketed in. In response to this, and as a means of replacing the numerous different classification systems, with a single controlled and unified classification system, the United Nations (UN) developed the current internationally agreed, standard scoring system, known as the Globally Harmonized System (GHS), also known as the "purple book" (UN, 2013). The GHS utilizes pictograms, signal words, hazard and precautionary statements, and safety data sheets according to standardized levels of physical, health and environmental hazards. The GHS is based upon averaged single tissue observations which can account for the reversibility of the observed chemical effects (Eskes et al., 2005). With regards to eye irritation, there are two primary categories. Substances which cause serious irreversible (up to 21 days) damage/destruction to the cornea, iris and/or conjunctiva are Category 1; substances which cause reversible (within 21 days) irritation including corneal opacity, iritis, redness or chemosis are Category 2. Category 2 chemicals can be split into two subcategories: 2A, irritating to eyes, chemicals which cause reversible irritation to eyes within 21 days; and 2B, mildly irritating to eyes, chemicals

which cause reversible irritation to eyes within 7 days. Non irritating chemicals are assigned a GHS No Category classification. The categories are assigned based on calculations of a mean score following observational grading at 24, 48 and 72 h post application of the test chemical. The GHS was adopted in 2002 and published in 2003 (Silk, 2003).

Despite the adoption of the GHS, Draize testing is often criticized due to its subjective and time consuming nature, lack of repeatability, variable estimates, insufficient relevance of test chemical application (Davila et al., 1998), high dosages (Curren and Harbell, 2002) and over-prediction of human responses (Jester et al., 2001), primarily due to interspecies differences. In addition, for most routine and acute toxicity tests, for example skin toxicity tests, there are standardized exposure times and/or delivery methods in place. This is not the case with eye irritation; liquids, pastes and solids all have different contact times with the eye, none of which are well defined (Prinsen, 2006). Draize testing also fails to elucidate the underpinning cellular and molecular mechanisms of toxicology. Since Draize assessments are based upon penlight or slit-lamp assessments, they provide very little information regarding the primary or secondary responses in the cornea, iris or conjunctiva (Maurer et al., 2002). Despite its "gold standard" status, Draize testing was never formally validated to any significant degree (Freeberg et al., 1986b). Since the anatomy of the rabbit eye differs from the human eye structurally, physiologically and biochemically, differences in sensitivity to irritants can occur. For example, in comparison to humans, rabbit corneas are thinner, have lower tear production, blinking frequency and ocular surface sensitivity (Huhtala et al., 2008). Rabbits have larger conjunctival sacs and a nictitating membrane (third eyelid), which may aid the removal of a test substance from the ocular surface (Calabrese, 1987).

There is almost no other field of science in which the fundamental experimental protocols have remained relatively unchanged for more than 40 years (Hartung, 2009), and yet consumers continually expect increased safety and information regarding their products. Worldwide, approximately £10 billion is spent on animal experimentation per annum, approximately £2 billion of which is on toxicological studies (Hartung, 2009). The cost associated with using, housing and maintaining colonies of live animals for toxicology testing of a single compound can exceed millions of pounds (Davila et al., 1998). Ethical (animal welfare), business (time and cost), scientific advances (reproducibility, mechanistic understanding) and legal concerns have all driven the demand for alternative, preferably animal-free testing platforms and protocols which are more precise and relevant to humans. There has been more focus on developing alternative testing techniques to Draize than all other *in vivo* toxicity tests combined (Huhtala et al., 2008). However, the development of alternative models has not advanced in a steady or continuous manner (Dholakiya and Barile, 2013), although the ban on animal testing for cosmetics use (Regulation (EC) N°. 1223/2009) has acted as a key driver for the development of alternative methods since this sector is constantly having to provide innovative and safe products. In Europe, with directive 2010/63/EU, there is a legal requirement to use alternatives where they exist. However, the reduction of animal use is primarily concentrated on toxicology studies since no government agency to date has eliminated animal use in basic biomedical research or pharmaceutical development.

3. Alternative *in vivo* tests

3.1. Low-volume eye-irritation test (LVET)

Low-volume eye-irritation tests (LVET) were developed in response to a recommendation from the National Research Council (NRC, 1977). LVET is a refinement of Draize testing developed by Griffith et al. (1980). The primary difference to the Draize test is that lower volumes of test substances (0.01 ml/0.01 g) (Lambert et al., 1993) are applied to the right-eye of the animal (Maurer et al., 2002), with no forced eyelid closure employed (ICCVAM, 2010b). Test substances are also only applied to the corneal surface and not the conjunctival sac. The test is believed to be less stressful to the tested animal (Jester et al., 2001). Pathological changes are characterized in the cornea, conjunctiva and iris/cilliary body (Maurer et al., 2002). Most LVET data is based upon surfactant-based mixtures or responses that are associated with mild irritation or non-irritants. This is due to the importance of surfactant use in cosmetic, pharmaceutical and household cleaning products (Davila et al., 1998). However, Gettings et al. (1996) investigated LVET in response to severe irritants and reported an under-prediction of results when compared to Draize data. Since Draize testing is often criticized for its over-prediction of human responses, it is arguable that LVET testing is more accurate (Freeberg et al., 1984, 1986a; Ghassemi et al., 1993; Roggeband et al., 2000). However, LVET is still criticized for its use of animals. In addition, should a negative irritancy result occur using a lower test volume, the standard procedure is to increase the concentration of the drug, effectively resorting back to Draize testing. The Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) recently evaluated the validity of LVET for the replacement of Draize testing. It was not considered to be a valid replacement nor recommend for prospective ocular safety testing (ICCVAM, 2010b). As a result, LVET has yet to be adopted by any regulatory agency as an alternative test. The reluctance to adopt LVET may be due to the fact that it does not offer the element of "exaggeration" present in Draize testing, that helps to assure public safety (Freeberg et al., 1986b; Ubels and Clousing, 2005). However, retrospective LVET data is still useful to weight-of-evidence approaches.

3.2. Human data

It has been suggested that the "gold standard" for eye irritation should be the human response (Bagley et al., 2006) and that ideally, a testing strategy to determine if a substance is harmful to humans would utilize an extremely high number of human subjects in order to faithfully represent human diversity. They would have to be unknowingly exposed to a substance under realistic conditions and the effects assessed (Hartung, 2009). However, such experimentation is both unrealistic and unethical. As a result, human study data and experiences of potential ocular hazards are only available from either accidental exposure or clinical studies. Unfortunately, accidental exposure data often does not realistically represent the most severe lesions since exposure is often brief due to immediate flushing of the eye. Furthermore, detailed studies by ophthalmologists and emergency room clinicians are often not collected or recorded. Clinical studies using undiluted products raise both scientific and ethical concerns, so experiments have to be carefully controlled. In clinical studies, the test material (often very small volumes and/or diluted) is usually applied to the upper or lower conjunctival sac, as opposed to the apex of the cornea as in *in vivo* rabbit studies (Freeberg et al., 1986b). This in itself raises concerns about the comparability of the outcomes. In addition, human testing often investigates the "sting" more so than irritation (Freeberg et al., 1986b). Studies performed in the 1980s compared results from hundreds of accidental human exposures with Draize and LVET tests (Freeberg et al., 1984, 1986a,b). In such a study using human volunteers, household substances commonly associated with accidental exposure (shampoo, hand soap, fabric softeners), exposure data was collected under known, controlled conditions to establish the relationship between *in vivo* animal tests and human exposure effects (Freeberg et al., 1986b). It was demonstrated that

Draize testing was a poor predictor of accidental human eye exposure, whereas LVET correlated well, although still over-predicted results.

Human studies are limited, and are usually comparing human responses with Draize or LVET, as proof-of-principle that LVET is more credible than Draize testing (Roggeband et al., 2000), and not as a comparison for the validation of alternative methods. A prevalent problem is that there is no human database for the development of the prediction models needed in validation studies, thus *in vitro* toxicity tests are still being compared to rabbit data (Bagley et al., 2006).

4. *Ex vivo* tests

4.1. Isolated/enucleated organ/organotypic methods

Ocular organotypic models are isolated systems that aim to maintain short-term normal physiological and biochemical function of the enucleated eye or cornea (Barile, 2010). The test material is often applied neat so is more relevant to industrial testing (Reader et al., 1990) and more faithfully represents accidental exposure. The protocols usually utilize opacitometric and spectroscopic methods for quantitative assessment of changes to the isolated cornea in response to a test material followed by histological analysis. Corneal opacity is also an *in vivo* corneal endpoint, although the data is observational, so often subjective. Corneal opacity acts as an indicator of protein denaturation, swelling, vacuolation and damage to the epithelium and corneal stroma (Barile, 2010). Fluorescein retention/leakage of the cornea is often used as a measure of permeability (Prinsen and Koëter, 1993), although *in vivo* the iris and the conjunctiva are also involved in ocular irritation, so corneal swelling and histological analysis are often included as additional endpoints in organotypic models (OECD, 2009a), often to distinguish "borderline" cases. Unlike *in vivo* testing, the quantitative assessment of corneal opacity and swelling provides more solid data, allowing for inter-laboratory variations to be easily determined. These measurements are then combined to derive an eye irritation classification or an *in vitro* irritancy score. Eye irritation is primarily determined by the extent of initial injury that correlates with the extent of cell death and ultimately the outcome of an irritant on an eye (Jester et al., 2001). Generally, slight irritants damage the superficial epithelium, mild irritants penetrate further to damage the stroma and severe irritants penetrate through the cornea and damage the endothelium (Jester et al., 2001) (Fig. 1).

Ocular organotypic models or enucleated eye tests (EET) were first introduced by Burton et al. (1981) using isolated rabbit eyes (IRE) from animals used for other research purposes, or those that had been sacrificed commercially as a food source (ICCVAM, 2010c). The IRE test, or rabbit enucleated eye test (REET) was originally

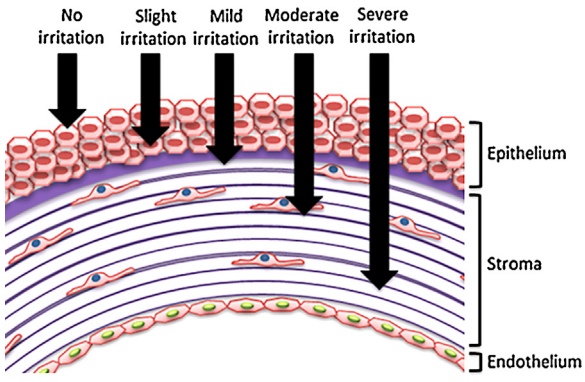

Fig. 1. Test substances that penetrate the furthest into the corneal tissue generally cause the most severe irritation.

developed to detect severe irritants that cause serious irreversible eye damage (Guo et al., 2012). Currently, the most commonly used test substances for IRE are active pharmaceutical ingredients, chemical/synthetic intermediates, cleaners, raw materials, soaps and detergents, solvents and surfactants (ICCVAM, 2010c). Lab-specific IRE protocols have developed over time, with variables including the evaluation of one to four different endpoints, differences in prediction models or classification systems, differences in the number of controls used and methodological variations (ICCVAM, 2010c). IRE has been extensively evaluated by international regulatory bodies including the European Commission/British Home Office (EC/HO), the Cosmetic, Toiletry and Perfumery Association (CTPA) and the Interagency Regulatory Alternatives Group (IRAG) (Guo et al., 2012). However, to date, the IRE protocol is not considered to be adequately validated for classification of ocular irritancy. Instead, it is advised that IRE is used for non-regulatory optimization studies to facilitate the collection of data to expand toxicology databases (ICCVAM, 2010c).

Slaughterhouse waste has been extensively investigated as an alternative tissue source (Prinsen, 1996) for EETs. Bovine or porcine corneas are often used (Reichl and Muller-Goymann, 2001), although chicken enucleated eye tests (CEET), also known as the isolated chicken eye (ICE) test are widely accepted to be a reliable and accurate slaughterhouse tissue for assessing the eye irritation potential of test materials (Prinsen, 1996). The ICE testing protocol (TG 438, (OECD, 2013b) is based upon the IRE model and was first described by Prinsen and Koëter (1993). The eyes are isolated from an intact chicken head and processed 2 h postmortem. The enucleated eye is then positioned in a clamp, with the cornea positioned vertically and transferred to a superfusion apparatus for examination of damage (Maurer et al., 2002) (Fig. 2i). Once approved, the eyes are equilibrated for up to 1 h (Fig. 2ii). Baseline thickness and opacity measurements are then recorded, before the eye is positioned horizontally and the test substance applied (0.03 ml liquid, 0.03 g solid) for 10 s (Fig. 2iii). The cornea is then rinsed with hypertonic saline (Fig. 2iv) before being returned to the superfusion chamber for analysis (Fig. 2v). Toxic effects are recorded by measuring changes in opacity, fluorescein retention, tissue thickness (swelling) and a macroscopic evaluation of changes to the surface of the tissue (OECD, 2013b).

A recent re-evaluation of ICE testing resulted in an endorsement for the test as being scientifically sound and that the test can be successfully used to identify substances that do not require classification (non-irritants, GHS No Category) as well as those deemed to cause serious irreversible eye damage (GHS Category 1). This guidance was adopted in 2009 (OECD, 2009a) and updated in 2013 (OECD, 2013b). Solids (soluble and insoluble), liquids, emulsions and gels can all be tested, although gases and aerosols have yet to be assessed and validated using this method. When used to identify GHS Category 1 chemicals, ICE has an overall accuracy of 86%, when used to identify GHS No Category chemicals ICE has an overall accuracy of 82% (OECD, 2013b). ICE is often used as a pre-screen for Draize testing; although despite promising outcomes the *in vivo* Draize testing results still overrule *ex vivo* results should discrepancies occur. Discrepancies are often associated with high false positive results for alcohols, and high false negative results for solids, surfactants and anti-fouling organic solvent containing paints (OECD, 2009a). ICE cannot be used to classify GHS Category 2, 2A or 2B chemicals, although to date, no *ex vivo* or *in vitro* test is capable of classifying chemicals in this category.

The Bovine Cornea Opacity Permeability (BCOP) assay was first developed by Gautheron et al. (1992) based on methods originally described by Muir (1984, 1985, 1987) and Tchao (1988). The intact corneas of healthy animals are held between O-rings mounted over a (posterior) chamber; an anterior chamber is positioned above the

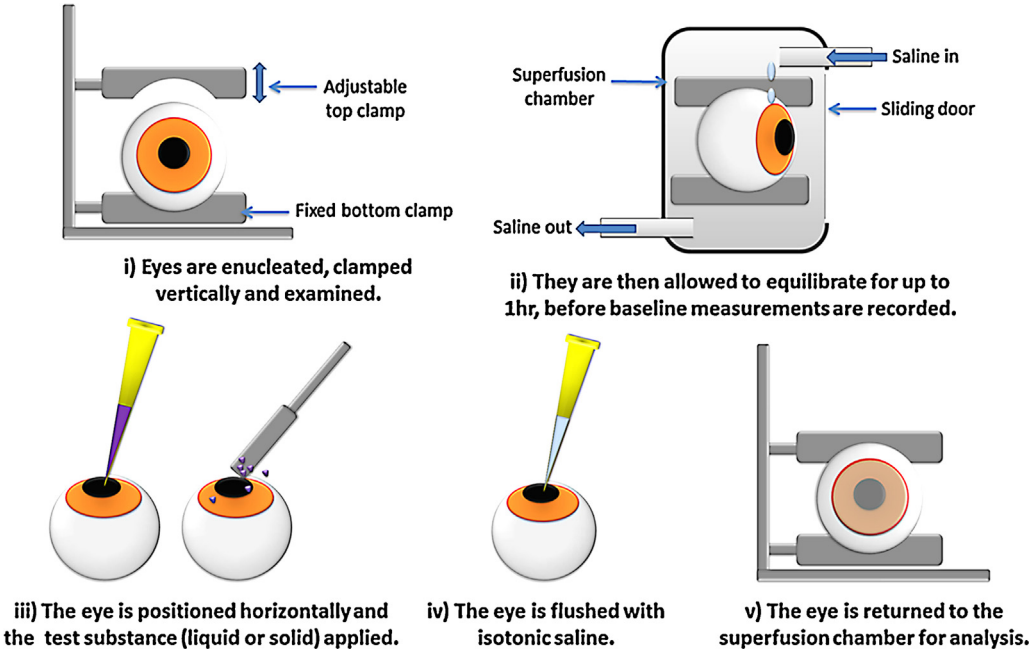

Fig. 2. Schematic representation of the chicken enucleated eye test (CEET), also known as the isolated chicken eye (ICE) protocol, which is based upon the isolated rabbit eye (IRE) protocol.

cornea, both of which are clamped together (Fig. 3). Each chamber has its own dosing hole which allows both the epithelium and endothelium to be treated independently. Currently, opacity is measured using an OP-KIT opacitometer, which provides a center-weighted reading of light transmission by measuring the changes in voltage when the transmission of white light alters as it passes through the cornea (Verstraelen et al., 2013). However, opacity readings can be underestimated as opaque areas tend to develop in spots in a non-homogeneous manner around the corneal periphery (Verstraelen et al., 2013). In response to this Van Goethem et al. (2010) developed a prototype laser light based opacitometer (PLLBOA), which has been further improved *via* the introduction of a camera with a speckle noise reducer, and an optimization of treatment conditions to more closely mimic the *in vivo* scenario (Jamur and Oliver, 2010; Verstraelen et al., 2013). A minimum of three eyes are used per test. Two different treatment protocols are

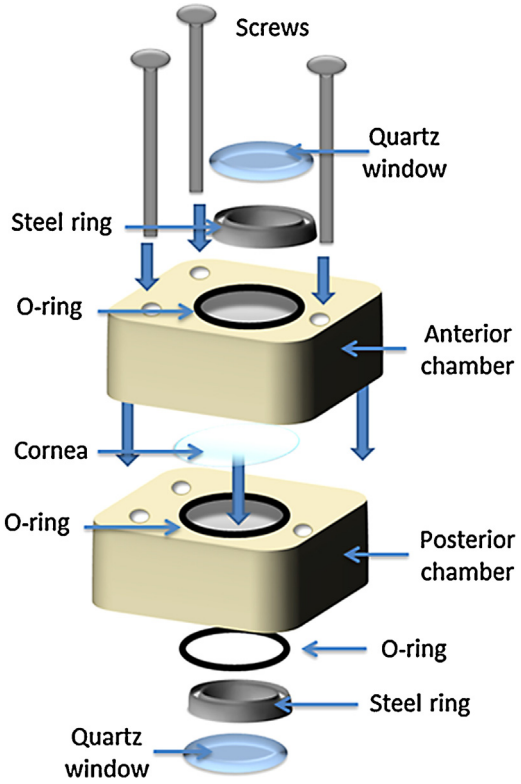

Fig. 3. Testing chamber utilized in Bovine Cornea Opacity Permeability (BCOP) and porcine cornea opacity (PCOP) testing.

used dependent upon whether the test material is a surfactant or not. An advantage of this assay is its speed, with results usually obtained within 24 h.

BCOP testing has been evaluated numerous times by ICCVAM, in conjunction with the European Union reference laboratory for alternatives to animal testing (EURL-ECVAM), formally known as the European Centre for the Validation of Alternative Methods (ECVAM) and the Japanese Centre for the Valuation of Alternative Methods (JaCVAM) regarding its suitability in identifying both substances that induce serious damage and those that are classified as non-irritants. It has been determined that BCOP is suitable and scientifically valid for both purposes (OECD, 2013a) and is routinely used by cosmetics and drug development companies for in-house testing of process intermediates (Eskes et al., 2005). Although it cannot be considered as a stand-alone test, BCOP received international acceptance in 2009 (OECD TG 437) which was then reviewed and updated in 2013 (OECD, 2013a). It is recommended for identifying severe irritants without further testing (OECD, 2009b) and has received endorsement for being a scientifically valid alternative test (OECD, 2013a). BCOP and has an overall accuracy of 79% when used to classify GHS Category 1 irritants, when compared to Draize testing (OECD, 2009b, 2013a). Loss of accuracy has been linked to high false positive rates for alcohols, ketones and solid test materials. When these are excluded, BCOP accuracy increases to 85%. However, since all alcohols and ketones are not over-predicted, they are not considered to be out of the applicability domain of the test. Solid materials often result in variable data and irrelevant results when using Draize testing (Prinsen, 2006) since solid materials can also cause mechanical damage. With regards to the classification of test materials that do not promote serious eye damage (GHS No Category), BCOP has an overall accuracy of 69%. BCOP does have a high false positive rate of 69% when compared to Draize data, but this value, although seemingly high, is not critical, since non-irritating chemicals which have a low *in vitro* irritancy score (IVIS) will be tested using another adequately validated *in vitro* test data, or as a last option *in vivo* rabbit testing (OECD, 2013a).

The porcine cornea opacity permeability (PCOP) assay uses porcine corneas, which can be considered as advantageous in comparison to bovine corneas since there are fewer concerns regarding encephalopathy diseases (Van den Berghe et al., 2005). Anatomically, it more accurately resembles the human cornea with regards to structure and thickness, and porcine corneas have been regularly used in ophthalmic research (Lynch and Ahearne, 2013). Since corneal size is different between pigs and cows, the holder utilized in PCOP studies has been slightly modified, with smaller diameters, and reduced volumes of test substances (Van den Berghe et al., 2005). Preliminary studies concerning PCOP revealed that it can be used to accurately predict eye irritation for liquid and water soluble substances (Van den Berghe et al., 2005). However, it has yet to be adopted by regulatory bodies that seem to favor BCOP.

Organotypic/enucleated models are borderline between *in vivo* and *in vitro* systems and are advantageous in that they have fewer ethical connotations (Luepke, 1985) with reduced costs. Although promising results have been obtained from EETs they all share the common problem that interspecies differences regarding anatomy and physiology are still present. Such differences produce discrepancies in permeation studies and toxicity tests (Reichl et al., 2004; Reichl and Muller-Goymann, 2003). EET models also lack, or do not consider conjunctival and irradial issues, inflammatory response elements and corneal recovery or reversibility of lesions (Guo et al., 2012). They also only account for corneal effects and cannot predict systemic effects of substances, such as the lethality of certain pesticides (OECD, 2009a). Furthermore they can only be used for relatively short-term assessment periods (4 h), and so are not suitable for testing substances that produce effects over extended time frames. However, such problems are associated with all *ex vivo* testing methods and protocols.

4.2. Non-ocular organotypic models

The chorioallantoic membrane vascular assay (CAMVA), also known as the Hen's egg test (HET), or Hühner-embryonen test on CAM (HET-CAM), or simply CAM assay was first proposed by Luepe and Kemper (Luepke, 1985; Luepke and Kemper, 1986). CAM is the vascularized respiratory membrane found within the membrane of a fertilized chicken egg, with a vasculature and inflammatory process similar to the conjunctival tissue of rabbit's eyes. The test is used to provide qualitative information on the potential effects occurring in the conjunctiva following exposure to a substance, whilst evaluation of coagulation can be used to reflect potential corneal damage (NICEATM, 2006). Although CAM models are usually classified alongside ICE, BCOP and IRE models, they differ in evaluation criteria used (Barile, 2010) since they have the addition of vasculature (Curren and Harbell, 2002). The general protocol involves exposing the CAM (Fig. 4i), the application of the test material to the surface (0.2–0.3 ml liquid, 0.1–0.3 g solid) (Fig. 4ii), followed by rinsing (Fig. 4iii) and observation of changes to the membrane morphology which are assessed and scored (Fig. 4iv). Incubation times, relative humidity, number of replicates, breed of hen (although white leghorn breed is most often used), criteria for egg selection (age/weight), egg rotation, method of opening the eggshell, volume/weight/concentration of test substance used (although usually applied neat), use of positive/negative controls and exposure times may vary dependent upon the protocol used. This inadvertently leads to problems regarding intra-laboratory reproducibility. Most protocols observe the time in seconds

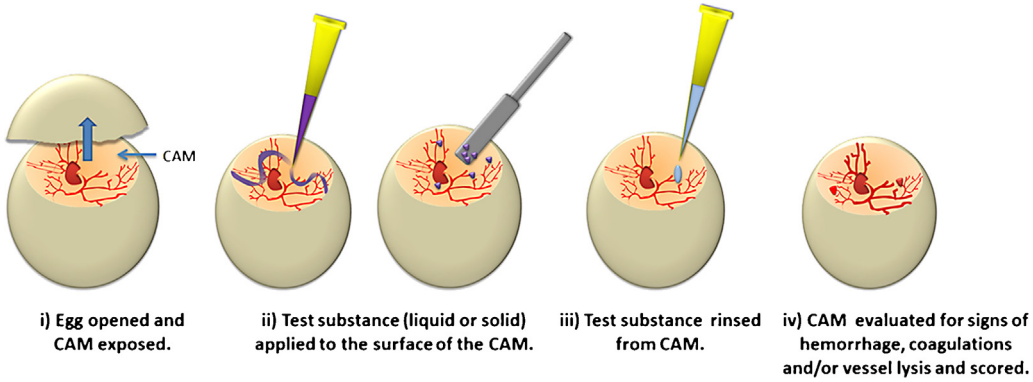

Fig. 4. Schematic representation of the chorioallantoic membrane vascular assay (CAMVA), also known as the Hen's egg test (HET), or Hühner-embryonen test on CAM (HET-CAM), or simply CAM assay.

whereby a substance causes hemorrhage, vasoconstriction and/or coagulation that is measured, scored and then categorized (Vinardell and Mitjans, 2008). Other endpoints include injection (mild hemorrhage), vasoconstriction, dilation and lysis (disintegration of vessels) (Gettings et al., 1996; Luepke, 1985; Luepke and Kemper, 1986; Macian et al., 1996; Spielmann, 1995; Sterzel et al., 1990). The irritation scoring varies dependent upon the classification system being used. The use of colored, turbid or substances that adhere to the CAM have been linked to compromised results since they impair visualization (NICEATM, 2006). The CAM assay has yet to receive international regulatory acceptance. Instead ICCVAM (2010a) recommends that the test is used for non-regulatory validation or optimization studies.

The slug mucosal irritation (SMI) assay was developed at the laboratory of pharmaceutical toxicology, Ghent University, Belgium to predict the mucosal irritancy potency of pharmaceutical formulations and ingredients (Adriaens et al., 2001, 2008; Adriaens and Remon, 1999). It uses the terrestrial slug *Arion lusitanicus*, which is considered to have limited sentience and so is not protected by legislation covering animal experiments (Adriaens and Remon, 1999). Slugs produce mucous and lose body weight when placed upon irritating surfaces. When tissue damage occurs the slug releases additional proteins and enzymes from its mucosal surface. Both of these factors allow for quantifiable endpoints, and for substances to be classified as non-irritating, irritating or severely irritating. In general, mild irritants cause an increase in mucous production, whereas severe irritants result in tissue damage and protein/enzyme release in addition to increased mucous production (Adriaens et al., 2008). In a previous study using 20 known reference chemicals it was shown that the SMI assay was a reliable and reproducible testing system (Adriaens et al., 2008). However the SMI assay failed to pass a formal validation study, so is currently only used as a pre-screen for simple toxicological endpoints.

5. *In vitro* tests

In vitro toxicity testing models and assays using cultured cells are advantageous compared to *in vivo* and *ex vivo* testing in that they are relatively inexpensive, simple, and quick to manufacture. This allows for replication and quantifiable data to be gathered, whilst also lending itself to automation. *In vitro* systems may also allow for a mechanistic understanding of toxicity at the cellular or molecular level (Davila et al., 1998). They are also capable of creating a broader range of toxicology testing capabilities such as multiple endpoints, concentrations, exposure methods and times to be better controlled and tailored accordingly.

One of the original alternative ocular irritation models was the EYTEX™ system which was developed, tested and evaluated in the 1990s (Courtellemont et al., 1999; Gordon et al., 1990; Matsukawa et al., 1999; Roy et al., 1994). Although EYTEX™ was unreliable at predicting ocular irritancy, primarily due to the lack of an appropriate prediction model; it did set the stage for the development of ocular toxicity models. The Ocular Irritection® assay is an updated protocol based upon the former EYTEX™ system (Eskes et al., 2005, 2014). The test is based upon the principle that eye irritation and corneal opacity caused by exposure to irritating chemicals alter the fundamental function of the proteins that make up the highly organized corneal tissue (Eskes et al., 2005). The assay is available as an off-the-shelf kit comprised of a macromolecular reagent of proteins, lipids, and low molecular weight proteins which when rehydrated form an ordered matrix similar to that of the native tissue, a membrane disc which allows for delivery of the test chemical, instrumentation and computer software. Test chemicals are gradually added using the defined membrane disc, resulting in turbidity of the matrix, due to the change in conformation and hydration (Eskes et al., 2005). Spectroscopic methods are used to measure the turbidity of the reagent at 405 nm. Prospective and retrospective validation studies have been performed to evaluate the suitability of the Ocular Irritection® assay for discriminating between chemicals that do not require classification from chemicals that do (Eskes et al., 2014). Limitations include limited usefulness with respect to intensely colored chemicals, underestimation of some cationic surfactants and overestimation of surfactant based formulations containing magnesium and multi-carboxylated carbohydrate chemicals (Eskes et al., 2005). Currently, the results of prospective and retrospective validation studies have been submitted for formal validation (Eskes et al., 2014).

5.1. Cytotoxicity assessment

Most *in vitro* ocular toxicity assays consist of a monolayer of cultured cells and a cytotoxicity assessment in response to a test material. In general, cytotoxicity measurements are quick, simple and inexpensive (Takahashi et al., 2008). Among the methods of assessing cytotoxicity are thymidine incorporation, Coomassie brilliant blue protein measurements, crystal violet and Lowry reagent, 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide assays (MTT assays), lactate dehydrogenase leakage (LDH), fluorescein leakage (FL) trypan blue exclusion, florescent staining with propidium iodide and neutral red uptake/release tests (Huhtala et al., 2008). Each of these methods has their advantages and limitations. In general, a combination of two or more of these methods is normally used to assess cytotoxicity.

Several assays combine cell staining with fluorescence or absorbance measurement to monitor changes in cell number and determine whether a substance is cytotoxic. Assays that use dyes such as trypan blue or propidium iodide are based on the concept that these dyes will be prevented from entering the cell unless there is disruption to the cells membrane (Strober, 2001). Hence healthy cells will remain unstained, while dead cells will stain positive. The amount of dye within a cell population can be measured and used to determine the percentage of cytotoxic cells. One limitation with this approach is that it only stains dead cells whilst dying or unhealthy cells may remain unstained. Alternatively a dye such as crystal violet can stain deoxyribonucleic acid (DNA) within a cell as shown (Fig. 5). In this assay the color absorbance of the stained cells can be measured at a wavelength of approximately 570 nm, which can then be used to assess the number of cells present (Gillies et al., 1986; Rothman, 1986). A reduction in cell number would indicate a cytotoxic effect. In the neutral red assay, lysosomes rather than DNA in healthy cells are stained positive. The dye can then be extracted and used to quantify the number of viable cells (Repetto et al., 2008). Fotakis and Timbrell (2006) found that the neutral red assay was more sensitive to cytotoxic effects on cells than several other assays tested.

In addition to staining, DNA can be quantified using other techniques. For example in a thymidine incorporation assay, 3H-thymidine (a radioactive nucleoside) is incorporated into newly synthesized DNA during mitosis. Inhabitation of thymidine incorporation would indicate cytotoxicity.

Protein assays have been used to determine cytotoxicity by measuring protein content within cells. A reduction in protein concentration would correspond to a decrease in the number of cells. Coomassie brilliant blue protein assays (also referred to as the Bradford assay) is a colorimetric protein assay that can be used to quantify cellular protein by measuring the color absorbance from stained cells. Similarly, the Lowry test measures the amount of cellular protein by reacting copper ions to amino acids in proteins under alkaline conditions and measuring a subsequent color change.

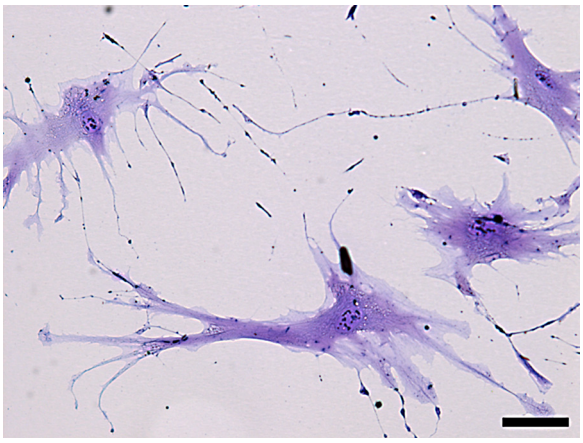

Fig. 5. Crystal violet staining of DNA in human adipose cell isolated from the vascular stromal fraction (scale bar 100 μm).

Enzymatic assays are among the most commonly used to assess cytotoxicity. LDH assays quantify the release of LDH following rupture of the cell membrane by using it to catalyze the conversion of lactate to pyruvate which can be measured colormetrically and used to quantify cell death. MTT assays measures the reduction of yellow MTT to purple formazan by mitochondrial succinate dehydrogenase. This change in color is measurable *via* spectrophotometry. As MTT reduction only occurs in metabolically active cells, the spectrophotometer reading can give an estimate of the number of viable cells present.

The short time exposure test (STE) is a relatively simple assay method that estimates cell cytotoxicity and viability using MTT (Kojima et al., 2013; Takahashi et al., 2008, 2011). It utilizes a confluent monolayer of rabbit corneal cell lines (Statens Seruminstitut rabbit corneal, SIRC) on a 96 well polycarbonate microplate (Takahashi et al., 2008). Test chemicals are dissolved or uniformly distributed in either physiological saline, 5% dimethyl sulfoxide (DMSO) in physiological saline, or mineral oils as test solvents, as opposed to culture medium which is often used in cytotoxic tests. This allows for water insoluble materials, acids and amides to be evaluated (Takahashi et al., 2008), which would otherwise have weakened effects when media is used as a solvent, due to the buffering effect that the media may have. As the name suggests, the exposure time to a given chemical is very short, it is only 5 min, compared to longer exposure times used in the FL assay (15 min) and the neutral red assay (1, 5 or 30 min) (Takahashi et al., 2008) for example. It is believed that that the short exposure is more similar to actual exposure conditions to a consumer product, whilst also providing fast results (Kojima et al., 2013; Takahashi et al., 2011). This also allows the STE to be used for high-throughput screening to evaluate many chemicals. Two different concentrations of the test material are evaluated, 5 and 0.5%, respectively. Post exposure cell viability is compared to a solvent control (relative viability) (OECD, 2014a; Takahashi et al., 2011). If the cell viability is $\leq$70% at both 0.5 and 5% concentration, then the chemical is classified as GHS Category 1. If cell viability if $\geq$70% at both concentrations then the chemical is classified as GHS No Category (OECD, 2014a). The STE was submitted to the OECD in 2011 as a method of high-throughput screening (Kojima et al., 2013) to evaluate minimal, moderate and severe eye irritation. The STE is currently under investigation *via* the OECD for regulatory acceptance as part of a tiered-testing strategy for either top–down or bottom–up approaches. It is recommended that STE is used for the identification of GHS Category 1, severe irritants and GHS No Category, non-irritants, although in both instances further testing is required to establish a definitive classification (OECD, 2014a). It is not recommended for the identification of GHS Category 2 (A or B) chemicals.

Penetration of a dye or reagent through a barrier of cells is another approach to assess cytotoxicity (Fig. 6). The FL assay (TG 460, (OECD, 2012c) can reveal the toxic effects of chemicals following a short exposure. A monolayer of Madin–Darby canine kidney (MDCK) cells are grown on permeable cell inserts. The test works by measuring the amount of fluorescein leakage through the cell monolayer which can be used to determine the integrity of the barrier formed by the cells. Cytotoxicity would result in an increase in the penetration of fluorescein through the monolayer. Increased *in vivo* permeability of the corneal epithelium correlates with the degree of inflammation and surface damage as eye irritation occurs. The volume of fluorescin leakage is measured spectrofluorometrically and compared to controls. The FL_{20}, which refers to the concentration that causes 20% FL relative to an untreated control is calculated and incorporated into a prediction model to identify irritation. FL is recommended for use in identifying severe, GHS Category 1 water soluble chemicals as part of a tiered-testing strategy (OECD, 2012c). Any chemical that is not predicted as severely irritating using FL would require further *in vitro* or *in vivo* methods, since it is not capable of distinguishing such chemicals. Another limitation is that FL cannot be used to classify strong acids and bases, cell fixatives and highly volatile chemicals since their modes of action cannot be measured using this mechanism. In addition, viscous and colored materials are not suited to this test (OECD, 2012c).

The Cytosensor Microphysiometer (CM) test method is a cytometric and cell based assay which utilizes a sub-confluent monolayer of mouse L3929 fibroblasts cultured on a transwell insert in a sensor chamber. Changes in acidity in response to an irritant are measured using a pH meter, ocular toxicity is evaluated by calculating the reduction in metabolic rate caused by the addition of a test chemical to the culture media compared to the basal metabolic state. CM has been recommended for the identification of GHS Category 1, severe irritants and GHS non irritants (Alépée et al., 2013; OECD, 2012a). The test is limited for use with test substances which do not settle or separate during analysis, primarily water soluble surfactants and surfactant-containing mixtures, but also some non-water-soluble solids, viscous chemicals or suspensions that maintain uniformity during analysis (OECD, 2012a). A draft OECD test guidance for CM is currently under review (OECD, 2012a).

Cell cultures using both target cells and non-target cells, usually expose cells to test materials that have been diluted in culture media (Reader et al., 1990), although both water soluble and insoluble materials can be assessed (Van Goethem et al., 2006). In general, cell culture methods are based upon long term cell survival, proliferation and function, including the release of

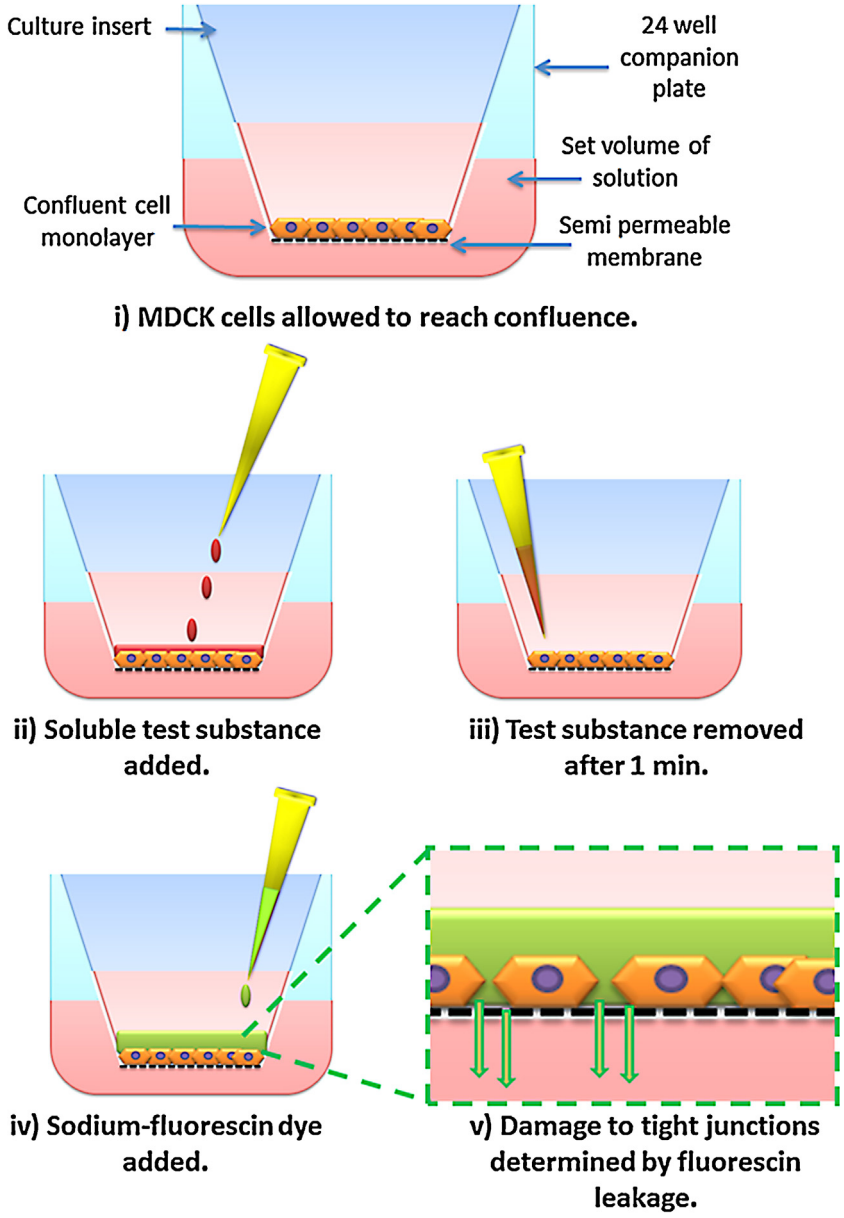

Fig. 6. Schematic diagram of the Fluorescein Leakage (FL) assay.

specific cytokines. Using permanent or immortalized cells lines is advantageous with regards to availability, reproducibility, ease of maintenance and ease of damage detection (Reader et al., 1990).

5.2. Corneal epithelial models

Several *in vitro* toxicity models have been developed using corneal epithelial cells. *In vivo* the epithelium is the outermost layer of the cornea that protects the underlying tissue by restricting foreign material from entering while still allowing gas and nutrient exchange to the underlying layers of the cornea. Thus it is the first point of contact for potentially hazardous materials. *In vitro* cultured epithelium is capable of retaining the *in vivo* repair mechanisms found in the native cornea (Davila et al., 1998), although these mechanisms are not always given the level of attention they deserve (Dholakiya and Barile, 2013). Epithelial models can be constructed from animal cells (commonly SIRC cells (Ubels and Clousing, 2005)) such as in the STE test, human epidermal cells, or human corneal cells, which are usually cultured in defined medium on cell culture membranes using air-lifting techniques (Alépée et al., 2013; Cotovio et al., 2007; Kaluzhny et al., 2011; Matsuda et al., 2009) to create a 3D stratified epithelium. Cytotoxicity following topical exposure is generally used as an endpoint (Curren and Harbell, 2002), and epithelial models have the potential to identify non-classified/non-irritating substances from mild irritants (Scott et al., 2010). Time-to-toxicity measurements (ET_{50}), which account for the time required for a 50% reduction in cell or tissue viability following exposure when compared to a negative control (Kaluzhny et al., 2011; Osborne et al., 1995), are often used as an endpoint. Although human primary epithelial cells have been investigated (Tripathi and Tripathi, 1988, 1989; Tripathi et al., 1989), their use is limited in toxicology models due to the lack of availability of human corneas and difficulties associated with expanding and passaging primary epithelial cells. Thus, rabbit corneal cells or mouse fibroblasts are often utilized as an alternative source.

Matsuda et al. (2009) cultured rabbit corneal epithelium (RCE) cells onto collagen hydrogels, which acts as a perabasal membrane. To validate the model 30 chemicals with known degrees of eye irritation (from Draize testing), ranging from non-irritating to severely irritating were tested. Inconsistencies occurred when

testing acids and alcohols, which was thought to be due to a pH dilution, the volatility of the alcohol, or a reaction with the buffer solution prior to testing (Matsuda et al., 2009).

The MatTek Corporation developed a commercially available 3D corneal epithelial model (OCL-200) based upon human derived epidermal keratinocytes from neonatal human foreskin (McLaughlin et al., 2009; Sheasgreen et al., 2009) grown on cell-culture inserts in serum-free media, to form a stratified, squamous epithelium, marketed as EpiOcular™. Test substances are directly applied to the models, and cytotoxicity is measured using MTT. Substances that cause the most rapid injury to cells generally have higher irritation potentials (Matsuda et al., 2009). The original protocol has since been developed into a single time-point protocol known as the EpiOcular™ eye irritation test (EIT) (Pfannenbecker et al., 2012). If the treated cells have viability greater than 60% post treatment then the test substance is classified as non-irritating. EpiOcular™ is currently used by numerous contract research laboratories, industrial cosmetic, personal care, and household chemical companies in place of Draize testing for product development.

SkinEthic Laboratories (Nice, France) developed a standardized 3D reconstituted human corneal epithelium (HCE) model from immortallized human corneal epithelial mucosa cells (Cotovio et al., 2007; Doucet et al., 2006; Van Goethem et al., 2006), that is structurally very similar to the corneal mucosa of the human eye. Substances are tested using two different exposure times: (i) short term, whereby the HCE is exposed to a chemical or substance for 10 min, (ii) long term, whereby the HCE is exposed for 60 min followed by 16 h incubation. If the treated HCE has viability greater than 50% post treatment then it is classified as non-irritating.

Both EpiOcular™ EIT and SkinEthic™ HCE models have undergone prospective validation by EURL-ECVAM and Cosmetics Europe to distinguish irritants (GHS Classification 1/Category 2) from non-irritants (GHS No Category) (Pfannenbecker et al., 2012; Zuang et al., 2013). Over 100 chemicals were tested and both methods showed high reproducibility (>90%) (Zuang et al., 2013). The EpiOcular™ EIT met all the predictive capacity acceptance criteria for the testing of liquids protocol, but not all of these criteria were met by the solids protocol nor by any of the SkinEthic™ HCE protocols (Zuang et al., 2013). The EpiOcular™ EIT solids protocol was further optimized for solids and further validation was conducted. At present the SkinEthic™ HCE model is still undergoing further optimisation for its solids protocol (Zuang et al., 2013). The final sensitivity of EpiOcular™ EIT was determined to be 96%, with specificity of 63% and accuracy of 80%, thus was considered valid for distinguishing non-irritants from irritants (OECD, 2014a). However, EpiOcular™ EIT is not intended to differentiate between GHS Category 1 (serious eye damage) and GHS Category 2 (eye irritation). A draft test guidance for EpiOcular™ EIT and performance standards has been delivered to the OECD (2014a), and the final test guidelines are expected to be adopted in 2015.

Recently Katoh et al. (2012, 2013) and Jung et al. (2011) developed the LabCyte CORNEA-MODEL (Japanese Tissue Engineering Co., Ltd., Japan) and MCTT-HCE model (MCTT, Seol, Korea), respectively. Unlike the commercially available EpiOcular™ EIT and SkinEthic™ HEC models, both utilize normal human corneal epithelial cells isolated from the human limbus of remaining corneal rim following transplantation that are cultured above and supported by cell feeder layers. They have been shown to express similar morphology and biomarker expression to the intact human corneal epithelium. Pre-validation studies have been performed for the LabCyte CORNEA-MODEL, to determine optimum treatment time, volume, post-incubation time and rinsing protocols (Jung et al., 2011). Although both MCTT-HCE and LabCyte CORNEA-MODEL have reported promising results with a high degree of accuracy, neither has yet to enter a formal validation assessment.

One of the limitations with epithelial models is that the initial enzymatic or mechanical dissociation of corneal tissue precluding cellular isolation itself may evoke traumatic stimuli that result in a diverse range of responses from the cell (Davila et al., 1998). This may include the disruption of cellular structural integrity, the release of inflammatory mediators or cell differentiation, all of which can lead to differences between *in vivo* and *in vitro* substance biokinetics (Davila et al., 1998). This needs to be considered when using any cells for *in vitro* toxicology testing. The use of immortalized epithelial cell lines does not always faithfully represent corneal cell behavior *in vivo,* since the immortalization process or subsequent culturing conditions alter expression patterns. For instance, cell lines do not express cytokeratins (CK) such as CK3, 7, 8, 18 and 19 (Huhtala et al., 2008). This may make the identification of specific biomarkers of toxicology somewhat more challenging.

Epithelial models are often fragile and have to be handled very carefully to avoid drying and damaging the tissues. Cell detachment in culture can lead to a misinterpretation of data dependent upon the experimental endpoint (Davila et al., 1998). They are also somewhat limited in that they only model the epithelial layer and so cannot be used to determine the possible effects of substances that *in vivo* penetrate the stroma and endothelium, or the reversibility of the irritation. They also do not account for the fact that some materials or chemicals may affect the various parts of the eye differently (Reader et al., 1990) and that cell–cell interactions, namely those between the epithelium and adjoining stroma are pivotal to corneal responses (McLaughlin et al., 2009; Wilson et al., 1999, 2014). *In vitro* cell based assays are also devoid of hormonal, immune and neural influences. Although this makes them simpler and easier to interpret, it can also be seen as a limitation since it does not account for the interactions that occur throughout the whole tissue, especially when considering the complexity in an organ as specialized as the eye (Barile, 2010).

5.3. Corneal equivalents

In response to the limitations incurred from using *in vitro* corneal epithelial models, more complex multicellular assays or corneal equivalents, termed as such due to their similarity to real corneas, have been under development in order to more accurately replicate the complexity and inherent characteristics of the native cornea. Both animal and human cells have been incorporated into corneal equivalents. Many studies have attempted to culture human primary cells under the premise that they will have a greater capacity for determining human irritancy (Zieske et al., 2004). However, problems associated with the isolation, growth, maintenance and differentiation of corneal cells has meant that many researchers choose to use transformed or immortalized human cell lines or animal cells, since they are easier to culture.

Griffith et al. (1999) produced the first working equivalent of a human cornea using immortalized human corneal cells. The model was originally developed to help to understand why corneas fail to heal properly after laser eye surgery. A collagen–chondroitin sulfate substrate cross-linked with glutaraldehyde was used as a tissue matrix. Initially a thin layer of endothelial cells was grown in a culture dish. Keratocytes and support proteins were added before finally adding the final epithelial layer. The gross morphology, transparency and histology were reported to be similar to that of a natural cornea. Tests performed using mild detergents determined that the construct had a similar gene expression and wound-healing response when compared to human eye-bank corneas, albeit more sensitive. The stromal matrix was later modified to allow for recovery mechanisms following exposure to chemical treatments (Doillon et al., 2003), and this was later followed by the introduction of nerve–target cell interactions (Suuronen et al.,

2004). Dorsal root ganglia isolated from chick embryos were utilized as a neural source, since optimal function, maintenance and wound healing of many tissues is dependent to some extent on peripheral sensory innervations (Suuronen et al., 2004). The innervated corneal constructs were reported to have lower cell death rates when exposed to test chemicals compared to non-innervated equivalents. This suggests that the presence of nerves protects the epithelium from chemical irritation and possibly explains why previous non-innervated corneal models have been deemed over-sensitive when used in toxicity studies. This model still requires further development since many of the functional properties of the nerves remain unclear. These types of models may demonstrate more promise for clinical development as cadaveric alternatives for corneal transplantation rather than as models for toxicological testing.

Reichl et al. (2005) manufactured a human corneal equivalent for *in vitro* drug permeation studies by culturing immortalized epithelial, endothelial and stromal cells in a collagen hydrogel matrix. Three reagents commonly used in ophthalmic drugs to treat glaucoma and inflammatory diseases were tested and permeation data obtained was compared with those from excised porcine cornea and a porcine cornea construct (Reichl et al., 2004; Reichl and Muller-Goymann, 2003). Porcine corneas were investigated due to their relatively similar anatomy and physiology to the human cornea. The human cornea construct had similar epithelial barrier properties to a native cornea with only small ultrastructural differences, possibly due to lack of tears and blinking. There was increased permeability in the corneal equivalents compared to the exercised porcine cornea for all reagents tested, although the differences were relatively minor. Unfortunately there was no data available to compare these corneal equivalents with an excised human cornea (as in the studies by Griffith et al. (1999).

In general, *in vitro* models that include a stromal layer have the greatest potential to distinguish severe/corrosive eye irritants from other classes (Scott et al., 2010). However recent validation studies have demonstrated that there is no single *in vitro* ocular irritation test, combination of tests, or testing strategies capable of completely replacing Draize testing (Huhtala et al., 2008) for predicting the response of the full range of irritation classes. This is partly due to a lack of understanding of the underlying cellular and molecular mechanisms of eye irritation (Matsuda et al., 2009; Maurer et al., 2002), a possible lack of innervation (Suuronen et al., 2004), difficulties associated when comparing *in vitro* data with historical animal data due to the subjective scoring systems used and the fact that *in vitro* systems only partially model *in vivo* tests, insufficient prediction models, inappropriate statistical analysis (Eskes et al., 2005) and an apparent reluctance of regulatory bodies to accept new *in vitro* corneal constructs.

The principle disadvantages of using multicellular *in vitro* models for toxicity assays, is that like epithelial based assays, they still lack the complexity of a complete organ (Becker et al., 2006). For example, the composition of the aqueous humor and tear fluid, or the mechanical stress of the eyelids and tear flow (Tegtmeyer et al., 2001), intrinsic clearing mechanisms (tearing and blinking) (Davila et al., 1998) are not taken into account. In a natural cornea all of these factors are important to protect the eye and are increased when exposed to irritation. *In vitro* false positive results can be attributed to the continuous contact with a test compound (Davila et al., 1998), thus the mechanisms that mimic tear production and blinking may need to be incorporated into *in vitro* toxicity models. Alternatively, *in vitro* assessment of the concentration in which a test substance is pharmacologically or toxicology active and relevant *in vivo* should be assessed (Davila et al., 1998) since the extent of the initial response is a pivotal mechanistic factor that determines the outcome of ocular irritation (Jester et al., 2001; Maurer et al., 2002).

5.4. *Future of* in vitro *alternatives*

It is unlikely that any single test, cell monolayer, three-dimensional epithelium, or multicellular corneal equivalent will be capable of mimicking the complexities and numerous physiological parameters of an *in vivo* system following exposure to a given substance (Borenfreund and Puerner, 1985; Pfannenbecker et al., 2012). In fact, having a "one-size fits all" approach has largely been abandoned, with the intention of many *in vitro* systems is to be utilized as part of an integrated testing strategy using either top–down or bottom–up tiered-testing approaches (Engelke et al., 2013; Scott et al., 2010). Top–down approaches are for the identification of severe irritants, bottom–up approaches are for the identification of non-irritating substances (Barile, 2010; Engelke et al., 2013). This results in two testing strategies to replace one *in vivo* test and as of yet, neither strategy can be used to distinguish between non-irritants and mild irritants and this remains an on-going challenge (Barile, 2010; Nóbrega et al., 2012; Scott et al., 2010). The commercial aims of *in vitro* testing are to be faster and cheaper, although currently, the costs are roughly on par with Draize testing. It is preferable that the testing procedures can be performed without the need for specialist training or expensive equipment (Dholakiya and Barile, 2013). From a corporate standpoint *in vitro* tests require the same level of investment as they are currently making using *in vivo* tests, so they either don't care, or fail to see the benefits in switching. A large factor that affects the decision making of corporate companies is that they are selling to a local market, not just countries within the EU. For developing or newly industrialized countries, Brazil, Russia, India, China and South Africa, the underlying challenge is getting them to understand the roles of *in vitro* tests, which is a continuing educational challenge. In order to overcome these issues, a re-evaluation of currently used *in vitro* tests may be required (Nóbrega et al., 2012).

In vitro assays and models provide useful data that complement *in vivo* studies allowing for significant reductions in the numbers of animals used. In order realize this, it must be ensured that clear endpoints correlate between *in vivo* and *in vitro* tests (Maurer et al., 2002). In general, *in vitro* tests are validated against the Draize test (Lenoir et al., 2011), with few actually investigating their predictability compared to humans. Despite the lack of formal validation, *in vitro* tests still are commonly used by industry. For example, industrial toxicologists often use *in vitro* protocols for prioritizing products and ingredients for further development (Curren and Harbell, 2002). However, use of the Draize test is still permitted worldwide, with the exception of the cosmetics section within Europe. Although *in vitro* alternatives tests are available, whether they are actually being used in practice is questionable. Every country has its own regulations and data requirements. The EU may be consolidated, but everywhere else is not and regulations have to be negotiated one by one – this is a very slow process, with no one country worse than the other. Regulations are aimed at protecting humans, and regulators focus on this, the culture of animal welfare is different in every country.

6. *In silico* models

In silico models are computer generated models that can play a useful role in predicting the ocular toxicity of a substance. *In silico* models utilize repositories of existing *in vitro* and *in vivo* toxicology data to predict the toxicity of samples. Quantitative structure–activity relationships (QSAR) are used to quantify the relationship between a sample's chemical structure and the biological effects that result from the same chemical (Simon-Hettich et al., 2006). QSAR is primarily based on the concept that the activity of molecules can be predicted from their structure and that these predictions are quantifiable (Valerio, 2011). The biological response

is incorporated into computer algorithms that are then used to generate predictive models. Once enough data for a specific toxicological endpoint is collected, evaluated and weighted, then a generalized relationship between the test substances and its biological activity can be defined (Simon-Hettich et al., 2006). Several different commercially available and freely available modeling software packages have been developed, the applicability of which have been previously evaluated in detail (Lo Piparo and Worth, 2010). This type of modeling is also dependent on the availability of suitable high quality databases, several of which have been previously discussed (Valerio, 2009).

The primary advantages in using *in silico* models to predict toxicity; other than the fact that they do not require the use of animals or animal tissues; is their speed and relative low cost. *In vitro* and *in vivo* toxicity models may take weeks or months to generate results at considerable expense while *in silico* models can generate results in minutes using just a computer and software. The continuing increase in computer processing speeds over recent years has enabled more sophisticated software to be developed. Among the limitations associated with *in silico* models are its reliance on high quality data. This can be a particular problem when compiling data from different laboratories that may have produced differing results. Since the models are reliant on data generated using animal models and cell based assays, the limitations associated with these, such as interspecies variations in toxicological response, still exist. Other limitations with *in silico* models have been described previously (Valerio, 2009). In general, *in silico* models tend to be more useful in predicting a specific endpoint rather than a broad range of toxicological effects that may be produced from a test substance (Nigsch et al., 2009) and they generally used with other test methods rather than exclusively by themselves.

7. Regulation and validation of ocular toxicity testing

Finding suitable, regulatory approved and validated alternatives to animal testing is a crucial aim of toxicological research (Alépée et al., 2013) with regulatory bodies keen to adopt the use of protocols that modify and reduce the number of animals used in ocular testing procedures. For alternative methods to be successfully incorporated into safety assessment procedures, they need to demonstrate that they can provide at least an equivalent or preferably superior level of protection to that obtained with current methods (Vinardell and Mitjans, 2008). *In vitro* and other alternative testing methods have a long history in corporate decision-making regarding chemical safety and product formulation. However, for many years, alternative testing strategies were neglected as a definitive testing strategy in the regulatory context (Stephens and Mak, 2013). Thus, corporations would often accompany alternative testing methods with more historical animal-based methods (Stephens and Mak, 2013). In order to move away from this status quo of toxicity testing, it is important to have an understanding of regulatory testing requirements and assessment and why they were developed (Fowle et al., 2013).

7.1. Regulation

Numerous regulatory authorities and systems exist worldwide for the assessment and classification of potentially hazardous substances. Their principal objective is to assess the hazardous potential of substances that may come into contact with the eye in order to supply regulations, guidelines and recommendations for their safe use. This offers consumers or the end user protection *via* the communication of hazardous information and protective measures (ICCVAM, 2010b; Wilhelmus, 2001) to prevent misapplication and to minimize accidental exposure. Regulatory assessment is based upon “informed decisions” that are not purely scientific in nature. They have to take into account congressional directives, legal precedent, benefit/cost considerations and public values (Fowle et al., 2013). This sometimes frustrates scientists, alternative-testing supporters and stakeholders alike, since “good science” does not always drive decision making (Fowle et al., 2013).

EURL-EVCAM aims to promote scientific and regulatory acceptance of non-animal tests. Similarly, ICCVAM is an interagency committee made up of 15 US Federal agencies including the Consumer Product Safety Commission (CPSC), National Institute for Occupational Safety and Health Administration (NIOSHA) and the FDA. ICCVAM aims to facilitate the development, validation and regulatory acceptance of new and revised regulatory test methods that reduce, refine and replace the use of animals. It was originally developed as a committee of the National Committee of Environmental Health Sciences (NIEHS) in 1997, and was made permanent in 2000 under NICEATM. Since then ICCVAM has contributed to 63 alternative testing methods, 38 of which do not require live animals, although not all of them are concerned with ocular tests.

Several directives restrict and even prohibit the use of animal testing, for example the Amendment of the Cosmetic European Directive (2003/15/EC) imposed a ban on the use of animals for the testing of cosmetics and their ingredients. However, until recently companies could still market products that had been animal tested outside of the EU. A new cosmetic regulation replaced the Cosmetics Directive in 2009 (Regulation (EC) No. 1223/2009) and since July 2013, cosmetics and cosmetic ingredients tested on animals can no longer be sold in Europe, even if they have been tested elsewhere. This has promoted considerable progress in replacing animal models for chemical toxicology (Alépée et al., 2013).

7.2. Validation and regulatory acceptance

EURL-ECVAM (formally ECVAM) and ICCVAM were formed in Europe and the US, respectively, to deal with the validation of alternative testing methods (Vinardell and Mitjans, 2008). Validation refers to the formal assessment, or rigorous set of policies that challenge the specific objectives of a test method or model with regard to its relevance and reliability. This in turn provides the foundation to facilitate regulatory adoption and acceptance (Corvi et al., 2006; Stephens and Mak, 2013). Relevance refers to the extent to which a test or model correctly predicts/measures the biological effect of interest; reliability is the degree to which the data in the protocol is reproducible within the guidelines or protocol of the method (Barile, 2010).

Most protocols undergo a pre-validation stage, designed to prepare a test model or assay for further progression into a formal validation study. These may involve intra-laboratory studies to address protocol optimization (Phase I), transferability (Phase II) and performance (Phase III) (Van Goethem et al., 2006), so that prior agreements can be made on detailed protocols that prepares and aids the test model or test in the formal validation process.

There are typically two types of validation study: prospective and retrospective (Kandárová and Letašiová, 2011) and a combination of these approaches are usually applied in the formal validation process (Hartung et al., 2004). Prospective studies involve the generation of new data, whilst retrospective studies re-assess existing data under standardized, controlled conditions.

ECVAM have proposed a modular validation assessment (Hartung et al., 2004), comprised of 7 modules aimed at determining the performance characteristics, advantages and limitations of a model or test for a specific purpose (Kandárová and Letašiová, 2011). The modules are: (i) test definition, where the scientific objective of the

model or test, a mechanistic basis, a specific protocol including all standard operating procedures with clearly defined endpoints, methods of results interpretation *via* prediction models and specific controls used must be clearly defined; (ii) intra-laboratory variability assessment, to determine potential variations in data incurred due to different operators carrying out the protocol within the same laboratory set-up. This assessment stage is usually not so problematic, since laboratories developing a model or test would usually abandon or modify an irreproducible protocol prior to assessment submission (Ubels and Clousing, 2005); (iii) transferability, to demonstrate that the test can be repeated in different laboratory set-ups. In the case of *in silico* models, this is the ability of different operators to reproduce the model definition and predictions, which is often dependent upon the strength of the explanatory documentation provided; (iv) inter-laboratory variability, whereby three to four laboratories are typically asked to test a defined number of substances using the assessed method or model to highlight discrepancies. This also adds further information to a test's predictive capacity; (v) predictive capacity, which utilizes a prediction model, to demonstrate how accurately a test can predict toxicity compared to a reference test (*i.e.*, Draize testing). Usually a defined number of substances in at least three different laboratories are assessed. Ironically, this stage of assessment can be hindered by the low reliability of Draize testing (Ubels and Clousing, 2005); (vi) applicability domain, which involves defining the purpose to which a test can be applied including endpoints, chemical classes, test material and physiochemical properties; (vii) performance standards, these need to be established for each test. However, if a similar, previously validated method or model exists, then the validation process is much faster (Hartung et al., 2004). The assessment of each module is led by a validation management group (VMP), who will then make recommendations to either ensue to peer review with a completed dossier of the information, or to collect additional data (IHCP, 2013). A test cannot proceed to peer review without a VMG recommendation. A formal regulatory validation can take more than five years to achieve (Sheasgreen et al., 2009) and may only then be considered for regulatory acceptance once achieved.

Regulatory acceptance is a formal recognition that indicates a test method or model may be used for a specific purpose. Acceptance is usually followed by a formal adoption by the EU and the OECD, and inclusion into the EU test method regulations and a publically available OECD test guideline (IHCP, 2013). The OECD continuously updates existing test guidelines and restructures draft proposals for future adoption (Barile, 2010), to encourage industries to use updated validated tests, whilst submitting data based upon them (Stephens and Mak, 2013).

Most assessments of validation and regulatory acceptance have occurred since 2000, following the establishment of vital alternative testing centers and the drive initiated European Cosmetic Directive (Stephens and Mak, 2013). However, the lack of human data has arguably led to delays in establishing the validity of alternative tests (Freeberg et al., 1986b). Currently only a limited number of ocular toxicity assays have undergone validation and regulatory acceptance. BCOP, ICE and FL have been accepted by ICCVAM, EURL-ECVAM and OECD for testing ocular corrosion and severe irritation. CM has also been accepted but is still awaiting final publication of OECD test guidelines. Dholakiya and Barile (2013) summarized the validation status of several *in vitro* ocular toxicity assays. Since that time a number of changes have been made to the validation status of these tests. For example, updated guidelines have been issued by the OECD for the BCOP (OECD, 2013b) and ICE tests (OECD, 2013a). For both tests changes have been made concerning the identification of chemical that do not require classification to UN GHS. In addition, clarifications have been provided on the applicability of BCOP for testing alcohols, ketones and solids and an updated list of chemicals to be used to demonstrate technical proficiency when applying the ICE test. A recent report by EURL-ECVAM on alternative methods (Zuang et al., 2013) also provided an update on the regulatory validation of several *in vitro* assays. To date EpiOcular EIT has passed initial validation studies (Zuang et al., 2013) and guidelines are currently being drafted by the OECD (2014b). The HET and IRE have been rejected by ICCVAM while the EURL-ECVAM has requested further optimization of the test protocol.

8. Conclusion

We have presented an overview of current practices in ocular toxicity testing. While progress has been made in developing a range of alternative techniques to *in vivo* testing, further progress is required to reduce the dependency of toxicity testing on live animals. Among the issues that need to be addressed by regulatory bodies is whether Draize testing should still be considered as the "gold standard" and whether results obtained from such testing used to validate or evaluate alternative tests. In order to advance alternative testing methodologies, there needs to be active engagement and dialog between scientific and regulatory communities.

Conflict of interest

The authors declare that there are no conflicts of interest.

Transparency document

The Transparency document associated with this article can be found in the online version.

Acknowledgment

Funding from EPSRC Engineering, Tissue Engineering and Regenerative Medicine Fellowship (E-TERM, Grant number: EP/1017801/1) and the University of Nottingham HERMES Fellowship (Grant number: 13b/I9) is gratefully acknowledged.

References

Adriaens, E., Remon, J.P., 1999. Gastropods as an evaluation tool for screening the irritating potency of absorption enhancers and drugs. Pharm. Res. 16, 1240–1244.
Adriaens, E., Dierckens, K., Bauters, T.G., Nelis, H.J., Van Goethem, F., Vanparys, P., Remon, J.P., 2001. The mucosal toxicity of different benzalkonium chloride analogues evaluated with an alternative test using slugs. Pharm. Res. 18, 937–942.
Adriaens, E., Bytheway, H., De Wever, B., Eschrich, D., Guest, R., Hansen, E., Vanparys, P., Schoeters, G., Warren, N., Weltens, R., 2008. Successful prevalidation of the slug mucosal irritation test to assess the eye irritation potency of chemicals. Toxicol. In Vitro 22, 1285–1296.
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, 1476–1488.
Bagley, D.M., Casterton, P.L., Dressler, W.E., Edelhauser, H.F., Kruszewski, F.H., McCulley, J.P., Nussenblatt, R.B., Osborne, R., Rothenstein, A., Stitzel, K.A., 2006. Proposed new classification scheme for chemical injury to the human eye. Regul. Toxicol. Pharm. 45, 206–213.
Barile, F.A., 2010. Validating and troubleshooting ocular in vitro toxicology tests. J. Pharm. Toxicol. Methods 61, 136–145.
Becker, R.A., Borgert, C.J., Webb, S., Ansell, J., Amundson, S., Portier, C.J., Goldberg, A., Bruner, L.H., Rowan, A., Curren, R.D., 2006. Report of an ISRTP Workshop: progress and barriers to incorporating alternative toxicological methods in the US. Regul. Toxicol. Pharm. 46, 18–22.
Borenfreund, E., Puerner, J.A., 1985. Toxicity determined in vitro by morphological alterations and neutral red absorption. Toxicol. Lett. 24, 119–124.
Burton, A.B.G., York, M., Lawrence, R.S., 1981. The in vitro assessment of severe eye irritants. Food Cosmet. Toxicol. 19, 471–480.
Calabrese, E.J., 1987. Principles of Animal Extrapolation. Wiley Interscience, New York.
Corvi, R., Ahr, H.-J., Albertini, S., Blakey, D.H., Clerici, L., Coecke, S., Douglas, G.R., Gribaldo, L., Groten, J.P., Haase, B., 2006. Meeting report: validation of

toxicogenomics-based test systems: ECVAM-ICCVAM/NICEATM considerations for regulatory use. Environ. Health Perspect. 114, 420–429.
Cotovio, J., Grandidier, M.-H., Lelièvre, D., Bremond, C., Flamand, N., Loisel-Joubert, S., Van Der Lee, A., Capallere, C., Meunier, J.-R., Leclaire, J., 2007. The use of the reconstructed human corneal model (HCE) to assess in vitro eye irritancy of chemicals. AATEX 14, 343–350.
Courtellemont, P., Pannetier, M., Biesse, J.P., Larnicol, M., Baret, J.P., Breda, B., 1999. Evaluation of the EYTEX(TM) system in the COLIPA eye irritation program. Toxicol. In Vitro 13, 295–304.
Curren, R.D., Harbell, J.W., 2002. Ocular safety: a silent (in vitro) success story. Altern. Lab. Anim. 30, 69–74.
Davila, J.C., Rodriguez, R.J., Melchert, R.B., Acosta Jr., S.L., 1998. Predictive value of in vitro model systems in toxicology. Annu. Rev. Pharmacol. Toxicol. 38, 63–96.
Dholakiya, L., Barile, F.A., 2013. Alternative methods for ocular toxicology testing: validation, applications and troubleshooting. Expert Opin. Drug Metab. Toxicol. 9, 699–712.
Doillon, C., Watsky, M., Hakim, M., Wang, J., Munger, R., Laycock, N., Osborne, R., Griffith, M., 2003. A collagen-based scaffold for a tissue engineered human cornea: physical and physiological properties. Int. J. Artif. Organs 26, 764–773.
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, 499–512.
Draize, J.H., Woodard, G., Calvery, H.O., 1944. Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J. Pharmacol. Exp. Ther. 82, 377–390.
Engelke, M., Zorn-Kruppa, M., Gabel, D., Reisinger, K., Rusche, B., Mewes, K., 2013. A human hemi-cornea model for eye irritation testing: quality control of production, reliability and predictive capacity. Toxicol. In Vitro 1, 458–468.
Eskes, C., Bessou, S., Bruner, L., Curren, R., Harbell, J., Jones, P., Kreiling, R., Liebsch, M., McNamee, P., Pape, W., 2005. Eye irritation. Altern. Lab. Anim. 33, 47.
Eskes, C., Hoffmann, S., Facchini, D., Ulmer, R., Wang, A., Flego, M., Vassallo, M., Bufo, M., van Vliet, E., d'Abroscat, F., 2014. Validation study on the Ocular Irritection® assay for eye irritation testing. Toxicol. In Vitro 28, 1046–1065.
Fotakis, G., Timbrell, J.A., 2006. In vitro cytotoxicity assays: comparison of LDH, neutral red, MTT and protein assay in hepatoma cell lines following exposure to cadmium chloride. Toxicol. Lett. 160, 171–177.
Fowle, J.R., Jacobs, A., Fitzpatrick, S., 2013. Regulatory testing to inform decisions: national and international requirements. In: Allen, D.G., Walters, M.D. (Eds.), Reducing, Refining and Replacing the Use of Animals in Toxicity Testing. RSC Publishing, Cambridge, UK, pp. 44–89.
Freeberg, F., Griffith, J., Bruce, R., Bay, P., 1984. Correlation of animal test methods with human experience for household products. Cutan. Ocul. Toxicol. 3, 53–64.
Freeberg, F., Hooker, D., Griffith, J., 1986a. Correlation of animal eye test data with human experience for household products: an update. Cutan. Ocul. Toxicol. 5, 115–123.
Freeberg, F., Nixon, G., Reer, P., Weaver, J., Bruce, R., Griffith, J., Sanders III, P., 1986b. Human and rabbit eye responses to chemical insult. Fundam. Appl. Toxicol. 7, 626–634.
Gautheron, P., Dukic, M., Alix, D., Sina, J.F., 1992. Bovine corneal opacity and permeability test: An in vitro assay of ocular irritancy. Fundam. Appl. Toxicol. 18, 442–449.
Gettings, S., Lordo, R., Demetrulias, J., Feder, P., Hintze, K., 1996. Comparison of low-volume, draize and in vitro eye irritation test data: I. Hydroalcoholic formulations. Food Chem. Toxicol. 34, 737–749.
Ghassemi, A., Sauers, L., Bruner, L., Reer, P., Hall, R., 1993. Demonstrating the human safety of a new household cleaning (HSC) product using alternatives to the Draize eye irritation test. Presentation made at the US Society of Toxicology meeting.
Gillies, R.J., Didier, N., Denton, M., 1986. Determination of cell number in monolayer cultures. Anal. Biochem. 159, 109–113.
Gordon, V.C., Kelly, C.P., Bergman, H.C., 1990. Applications of the EYTEX method. Toxicol. In Vitro 314–317.
Griffith, J.F., Nixon, G.A., Bruce, R.D., Reer, P.J., Bannan, E.A., 1980. Dose–response studies with chemical irritants in the albino rabbit eye as a basis for selecting optimum testing conditions for predicting hazard to the human eye. Toxicol. Appl. Pharmacol. 55, 501–513.
Griffith, M., Osborne, R., Munger, R., Xiong, X.J., Doillon, C.J., Laycock, N.L.C., Hakim, M., Song, Y., Watsky, M.A., 1999. Functional human corneal equivalents constructed from cell lines. Science 286, 2169–2172.
Guo, X., Yang, X.F., Yang, Y., Hans, R., Cai, J.H., Xue, J.Y., Tan, X.H., Xie, X.P., Xiong, X.K., Huang, J.M., 2012. Prediction of ocular irritancy of 26 chemicals and 26 cosmetic products with isolated rabbit eye (IRE) test. Biomed. Environ. Sci 25, 359–366.
Hartung, T., 2009. Toxicology for the twenty-first century. Nature 460, 208–212.
Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Roi, A.J., 2004. A modular approach to the ECVAM principles on test validity. Altern. Lab. Anim. 32, 467–472.
Huhtala, A., Salminen, L., Tähti, H., Uusitalo, H., 2008. Corneal models for the toxicity testing of drugs and drug releasing materials. In: Ashammakhi, N. (Ed.), Topics in Multifunctional Biomaterials and Devices. Biomaterials and Tissue Engineering Group. http://www.oulu.fi/spareparts/ebook_topics_multifunctional/index.html.
ICCVAM, 2010a. National Institute of Environmental Heath Sciences. ICCVAM-Recommended Test Method Protocol: Hen's Egg Test – Chorioallantoic Membrane (HET–CAM) Test Method. National Institute of Environmental Heath Sciences, Triangle Park, NC NIH Publication.
ICCVAM, 2010b. In: D.a.M.P. International Association for Soaps (Ed.), ICCVAM LVET Test Method Evaluation Report: Recommendation to Discontinue Use of the Low Volume Eye Test for Ocular Safety Testing. National Institute of Environmental Heath Sciences, Triangle Park, NC NIH Publication.
ICCVAM, 2010c. In: N.I.o.E.H. Sciences (Ed.), ICCVAM Recommended Test Method Protocol Isolated Rabbit Eye Test Method. National Institute of Environmental Heath Sciences, Triangle Park, NC, pp. 1–15 NIH Publication.
IHCP, 2013. Validation and Regulatory Acceptance. JRC Publications. https://eurl-ecvam.jrc.ec.europa.eu/validation-regulatory-acceptance.
Jamur, M.C., Oliver, C., 2010. Cell fixatives for immunostaining. Methods Mol. Biol. (Clifton, N. J.) 588, 55–61.
Jester, J.V., Li, L., Molai, A., Maurer, J.K., 2001. Extent of initial corneal injury as a basis for alternative eye irritation tests. Toxicol. In Vitro 15, 115–130.
Jung, K.-M., Lee, S.-H., Ryu, Y.-H., Jang, W.-H., Jung, H.-S., Han, J.-H., Seok, S.-H., Park, J.-H., Son, Y., Park, Y.-H., 2011. A new 3D reconstituted human corneal epithelium model as an alternative method for the eye irritation test. Toxicol. In Vitro 25, 403–410.
Kaluzhny, Y., Kandárová, H., Hayden, P., Kubilus, J., d'Argembeau-Thornton, L., Klausner, M., 2011. Development of the EpiOcular™ 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. 39, 339.
Kandárová, H., Letašiová, S., 2011. Alternative methods in toxicology: pre-validated and validated methods. Interdiscip. Toxicol. 4, 107–113.
Katoh, M., Uemura, N., Hamajima, F., Ogasawara, T., Hata, K., 2012. Morphological characterization of a reconstructed human corneal epithelial model (LabCyte CORNEA-MODEL) as an alternative to the Draize eye test for the assessment of eye irritation. AATEX 17, 2–8.
Katoh, M., Hamajima, F., Ogasawara, T., Hata, K.-i, 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, 2184–2192.
Kojima, H., Hayashi, K., Sakaguchi, H., Omori, T., Otoizumi, T., Sozu, T., Kuwahara, H., Hayashi, T., Sakaguchi, M., Toyoda, A., 2013. Second-phase validation study of short time exposure test for assessment of eye irritation potency of chemicals. Toxicol. In Vitro 27, 1855–1869.
Lambert, L.A., Chambers, W.A., Green, S., Gupta, K.C., Hill, R.N., Hurley, P.M., Lee, C.C., Lee, J.K., Liu, P.T., Lowther, D.K., Roberts, C.D., Seabaugh, V.M., Springer, J.A., Wilcox, N.L., 1993. The use of low-volume dosing in the eye irritation test. Food Chem. Toxicol. 31, 99–103.
Lenoir, J., Claerhout, I., Kestelyn, P., Klomp, A., Remon, J.-P., Adriaens, E., 2011. The slug mucosal irritation (SMI) assay: development of a screening tool for the evaluation of ocular discomfort caused by shampoos. Toxicol. In Vitro 25, 1919–1925.
Lo Piparo, E., Worth, A., 2010. Review of QSAR Models and Software Tools for Predicting Developmental and Reproductive Toxicity. JRC Scientific and Technical Reports EUR, Publications Office of the European Union, Joint Research Centre Publications Repository, Luxemburg.
Luepke, N., 1985. Hen's egg chorioallantoic membrane test for irritation potential. Food Chem. Toxicol. 23, 287–291.
Luepke, N., Kemper, F., 1986. The HET–CAM test: an alternative to the Draize eye test. Food Chem. Toxicol. 24, 495–496.
Lynch, A.P., Ahearne, M., 2013. Strategies for developing decellularized corneal scaffolds. Exp. Eye Res. 108, 42–47.
Macian, M., Seguer, J., Infante, M.R., Selve, C., Vinardell, M.P., 1996. Preliminary studies of the toxic effects of non-ionic surfactants derived from lysine. Toxicology 106, 1–9.
Matsuda, S., Hisama, M., Shibayama, H., Itou, N., Iwaki, M., 2009. Application of the reconstructed rabbit corneal epithelium model to assess the in vitro eye irritancy test of chemicals. Yakugaku Zasshi 129, 1113–1120.
Matsukawa, K., Masuda, K., Kakishima, H., Suzuki, K., Nakagawa, Y., Matsushige, C., Imanishi, Y., Nakamura, T., Mizutani, A., Watanabe, R., Shingai, T., Kaneko, T., Hirose, A., Ohno, Y., 1999. Interlaboratory validation of the in vitro eye irritation tests for cosmetic ingredients. (11) EYTEX(TM). Toxicol. In Vitro 13, 209–217.
Maurer, J.K., Parker, R.D., Jester, J.V., 2002. Extent of initial corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. Regul. Toxicol. Pharmacol. 36, 106–117.
McLaughlin, C.R., Tsai, R.J.F., Latorre, M.A., Griffith, M., 2009. Bioengineered corneas for transplantation and in vitro toxicology. Front Biosci. 14, 3326–3337.
Muir, C., 1984. A simple method to assess surfactant-induced bovine corneal opacity in vitro: preliminary findings. Toxicol. Lett. 22, 199–203.
Muir, C., 1985. Opacity of bovine cornea in vitro induced by surfactants and industrial chemicals compared with ocular irritancy in vivo. Toxicol. Lett. 24, 157–162.
Muir, C., 1987. Surfactant-induced opacity of bovine isolated cornea: an epithelial phenomenon? Toxicol. Lett. 38, 51–54.
NICEATM, 2006. In: D.o.H.a.H. Services (Ed.), Current Status of in Vitro Test Methods for Identifying Ocular Corrosives and Severe Irritants: Hen's Egg Test–Chorioallantoic Membrane Test Method. ICCVAM Publication, NIEHS Publication, North Carolina.

Nigsch, F., Macaluso, N.J., Mitchell, J.B., Zmuidinavicius, D., 2009. Computational toxicology: an overview of the sources of data and of modelling methods. Expert Opin. Drug Metab. Toxicol. 5, 1–14.
Nóbrega, A.M.d., Alves, E.N., Presgrave, R.d.F., Costa, R.N., Delgado, I.F., 2012. Determination of eye irritation potential of low-irritant products: comparison of in vitro results with the in vivo draize rabbit test. Braz. Arch. Biol. Technol. 55, 381–388.
NRC, 1977. Principles and Procedures for Evaluating the Toxicity of Household Substances. National Academy of Sciences Publication, Washington, DC.
OECD, 2009a. Test No. 438: isolated chicken eye test method for identifying ocular corrosives and severe irritants. OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris.
OECD, 2009b. Test number 437: bovine corneal opacity and permeability test metjod for identifying i) chemicals induscing serious eye damage and ii) chemicals not requiring classification for eye irritation and serious eye damage. OECD Guidelines for the Testing of Chemicals, Section 4. OECD Publishing, Paris.
OECD, 2012a. Draft OECD Guideline for the Testing of Chemicals: The Cytosensor Microphysiometer Test Method: An in Vitro Method for Identifying Ocular 3corrosive and Severe Irritant Chemicals as Well as Chemicals Not Classified as Ocular Irritants. OECD Publishing, Paris.
OECD, 2012b. Test No. 405: Acute Eye Irritation/Corrosion. OECD Publishing, Paris.
OECD, 2012c. Test No. 460: Fluorescein Leakage Test Method for Identifying Ocular Corrosives and Severe Irritants. OECD Publishing, Paris.
OECD, 2013a. In: Guidleines, O. (Ed.), 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 Publishing, Paris.
OECD, 2013b. 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 Publishing, Paris.
OECD, 2014. Draft Guideline for the Testing of Chemicals. The 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, Paris.
OECD, 2014. OECD Guideline For The Testing Of Chemicals Draft Proposal For A New Test Guideline Reconstructed Human Cornea-like Epithelium (RhCE) Test Method for Identifying Chemicals Not Requiring Classification and Labelling for Eye Irritation or Serious Eye Damage, Paris.
Osborne, R., Perkins, M.A., Roberts, D.A., 1995. Development and intralaboratory evaluation of an in vitro human cell-based test to aid ocular irritancy assessments. Toxicol. Sci. 28, 139–153.
Pfannenbecker, U., Bessou-Touya, S., Faller, C., Harbell, J., Jacob, T., Raabe, H., Tailhardat, M., Alépée, N., De Smedt, A., De Wever, B., 2012. Cosmetics Europe multi-laboratory pre-validation of the EpiOcular™ reconstituted human tissue test method for the prediction of eye irritation. Toxicol. In Vitro 27, 619–626.
Prinsen, M.K., 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, 291–296.
Prinsen, M.K., 2006. The Draize eye test and in vitro alternatives; a left-handed marriage? Toxicol. In vitro 20, 78–81.
Prinsen, M., Koëter, H., 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, 69–76.
Reader, S.J., Blackwell, V., O'Hara, R., Clothier, R.H., Griffin, G., Balls, M., 1990. Neutral red release from pre-loaded cells as an in vitro approach to testing for eye irritancy potential. Toxicol. In Vitro 4, 264–266.
Reichl, S., Muller-Goymann, C.C., 2001. Development of an organotypical cornea construct as an in vitro model for permeation studies. Ophthalmologe 98, 853–858.
Reichl, S., Muller-Goymann, C.C., 2003. The use of a porcine organotypic cornea construct for permeation studies from formulations containing befunolol hydrochloride. Int. J. Pharm. 250, 191–201.
Reichl, S., Bednarz, J., Muller-Goymann, C.C., 2004. Human corneal equivalent as cell culture model for in vitro drug permeation studies. Br. J. Ophthalmol. 88, 560–565.
Reichl, S., Dohring, S., Bednarz, J., Muller-Goymann, C.C., 2005. Human cornea construct HCC – an alternative for in vitro permeation studies? A comparison with human donor corneas. Eur. J. Pharm. Biopharm. 60, 305–308.
Repetto, G., del Peso, A., Zurita, J.L., 2008. Neutral red uptake assay for the estimation of cell viability/cytotoxicity. Nat. Protoc. 3, 1125–1131.
Roggeband, R., York, M., Pericoi, M., Braun, W., 2000. Eye irritation responses in rabbit and man after single applications of equal volumes of undiluted model liquid detergent products. Food Chem. Toxicol. 38, 727–734.
Rothman, S.W., 1986. Technique for measuring 50% end points in cytotoxicity assays for Clostridium difficile toxins. J. Clin. Pathol. 39, 672–676.
Roy, T.A., Saladdin, K.A., Mackerer, C.R., 1994. Evaluation of the EYTEX system as a screen for eye irritancy of petroleum products. Toxicol. In Vitro 8, 797–798.
Russell, W.M.S., Burch, R.L., Hume, C.W., 1959. The Principles of Humane Experimental Technique. Methuen, London, UK.
Scott, L., Eskes, C., Hoffmann, S., Adriaens, E., Alepée, N., Bufo, M., Clothier, R., Facchini, D., Faller, C., Guest, R., 2010. A proposed eye irritation testing strategy to reduce and replace in vivo studies using bottom–up and top–down approaches. Toxicol. In Vitro 24, 1–9.
Shaw, A., Balls, M., Clothier, R., Bateman, N., 1991. Predicting ocular irritancy and recovery from injury using Madin–Darby canine kidney cells. Toxicol. In Vitro 5, 569–571.
Sheasgreen, J., Klausner, M., Kandarova, H., Ingalls, D., 2009. The MatTek Story – how the three Rs principles led to 3-D tissue success!. Altern. Lab. Anim. 37, 611–622.
Silk, J.C., 2003. Development of a Globally Harmonized System for hazard communication. Int. J. Hyg. Environ. Health 206, 447–452.
Simon-Hettich, B., Rothfuss, A., Steger-Hartmann, T., 2006. Use of computer-assisted prediction of toxic effects of chemical substances. Toxicology 224, 156–162.
Spielmann, H., 1995. HET-CAM test. In: O'Hare, S., Atterwil, C.K. (Eds.), In Vitro Toxicity Testing Protocols Methods in Molecular Biology. Humana Press, Totowa, NJ, pp. 199–204.
Stephens, M.L., Mak, N.S., 2013. History of the 3Rs in toxicity testing: from Russell and Burch to 21st century toxicology. In: Allen, D., Walters, M. (Eds.), Reducing, Refining and Replacing the Use of Animals in Toxicity Testing. RSC, Publishing, Cambridge, UK, pp. 1–39.
Sterzel, W., Bartnik, F., Matthies, W., Kästner, W., Künstler, K., 1990. Comparison of two in vitro and two in vivo methods for the measurement of irritancy. Toxicol. In Vitro 4, 698–701.
Strober, W., 2001. Trypan blue exclusion test of cell viability. In: Current protocols in Immunology. John Wiley & Sons, Inc., Bethesda, MD pp. A.3B.1–A.3B.2.
Suuronen, E.J., McLaughlin, C.R., Stys, P.K., Nakamura, M., Munger, R., Griffith, M., 2004. Functional innervation in tissue engineered models for in vitro study and testing purposes. Toxicol. Sci. 82, 525–533.
Takahashi, Y., Koike, M., Honda, H., Ito, Y., Sakaguchi, H., Suzuki, H., Nishiyama, N., 2008. Development of the short time exposure (STE) test: an in vitro eye irritation test using SIRC cells. Toxicol. In Vitro 22, 760–770.
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, 1425–1434.
Tchao, R., 1988. Trans-epithelial permeability of fluorescein in vitro as an assay to determine eye irritants. In: Goldbery, A. (Ed.), Alternative Methods in Toxicology. Mary Ann Liebert, New York, pp. 271–283.
Tegtmeyer, S., Papantoniou, I., Muller-Goymann, C.C., 2001. Reconstruction of an in vitro cornea and its use for drug permeation studies from different formulations containing pilocarpine hydrochloride. Eur. J. Pharm. Biopharm. 51, 119–125.
Tripathi, B., Tripathi, R., 1988. Cytotoxic effects of benzalkonium chloride and chlorobutanol on human corneal epithelial cells in vitro. Lens Eye Toxic. Res. 6, 395–403.
Tripathi, B.J., Tripathi, R.C., 1989. Hydrogen peroxide damage to human corneal epithelial cells in vitro: implications for contact lens disinfection systems. Arch. Ophthalmol. 107, 1516.
Tripathi, B., Tripathi, R., Millard, C., Borisuth, N., 1989. Cytotoxicity of hydrogen peroxide to human corneal epithelium in vitro and its clinical implications. Lens Eye Toxic. Res. 7, 385–401.
Ubels, J.L., Clousing, D.P., 2005. In vitro alternatives to the use of animals in ocular toxicology testing. Ocul. Surf. 3, 126–142.
UN, 2013. Globally Harmonized System for the Classification and Labelling of Chemicals (GHS), New York and Geneva, pp. 536.
Valerio Jr., L.G., 2009. In silico toxicology for the pharmaceutical sciences. Toxicol. Appl. Pharmacol. 241, 356–370.
Valerio Jr., L.G., 2011. In silico toxicology models and databases as FDA Critical Path Initiative toolkits. Hum. Genom. 5, 200–207.
Van den Berghe, C., Guillet, M., Compan, D., 2005. Performance of porcine corneal opacity and permeability assay to predict eye irritation for water-soluble cosmetic ingredients. Toxicol. In Vitro 19, 823–830.
Van Goethem, F., Adriaens, E., Alepee, N., Straube, F., De Wever, B., Cappadoro, M., Catoire, S., Hansen, E., Wolf, A., Vanparys, P., 2006. Prevalidation of a new in vitro reconstituted human cornea model to assess the eye irritating potential of chemicals. Toxicol. In Vitro 20, 1–17.
Van Goethem, F., Hansen, E., Sysmans, M., De Smedt, A., Vanparys, P., Van Gompel, J., 2010. Development of a new opacitometer for the bovine corneal opacity and permeability (BCOP) assay. Toxicol. In Vitro 24, 1854–1861.
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, 1298–1311.
Vinardell, M., Mitjans, M., 2008. Alternative methods for eye and skin irritation tests: an overview. J. Pharm. Sci. 97, 46–59.
Wilhelmus, K.R., 2001. The Draize eye test. Surv. Ophthalmol. 45, 493–515.
Wilson, S.E., Liu, J.J., Mohan, R.R., 1999. Stromal–epithelial interactions in the cornea. Prog. Retin. Eye Res. 18, 293–309.
Wilson, S.L., Yang, Y., El Haj, A.J., 2014. Corneal stromal cell plasticity: in vitro regulation of cell phenotype through cell–cell interactions in a three-dimensional model. Tissue Eng. Part A 20, 225–238.
Zieske, J.D., Chung, E.-H., Xiaoqing, G., Hutcheon, A.E., 2004. Human corneal organotypic cultures. J. Toxicol. Cutan. Ocul. Toxicol. 23, 19–28.
Zuang, V., Schäffer, M., Tuomainen, A.M., Amcoff, P., Bernasconi, C., Bremer, S., Casati, S., Castello, P., Coecke, S., Corvi, R., Griesinger, C., Janusch Roi, A., Kirmizidis, G., Prieto, P., Worth, A., Munn, S., Berggren, E., Whelan, M., 2013. In: European Commission, J.R.C., Institute for Health and Consumer Protection (Eds.), EURL ECVAM Progress Report on the Development, Validation and Regulatory Acceptance of Alternative Methods (2010-2013). Prepared in the Framework of Directive 76/768/EEC and Regulation (EC) No 1223/2009 on Cosmetic Products. Publications Office of the European Union, Luxembourg, pp. 1–63.