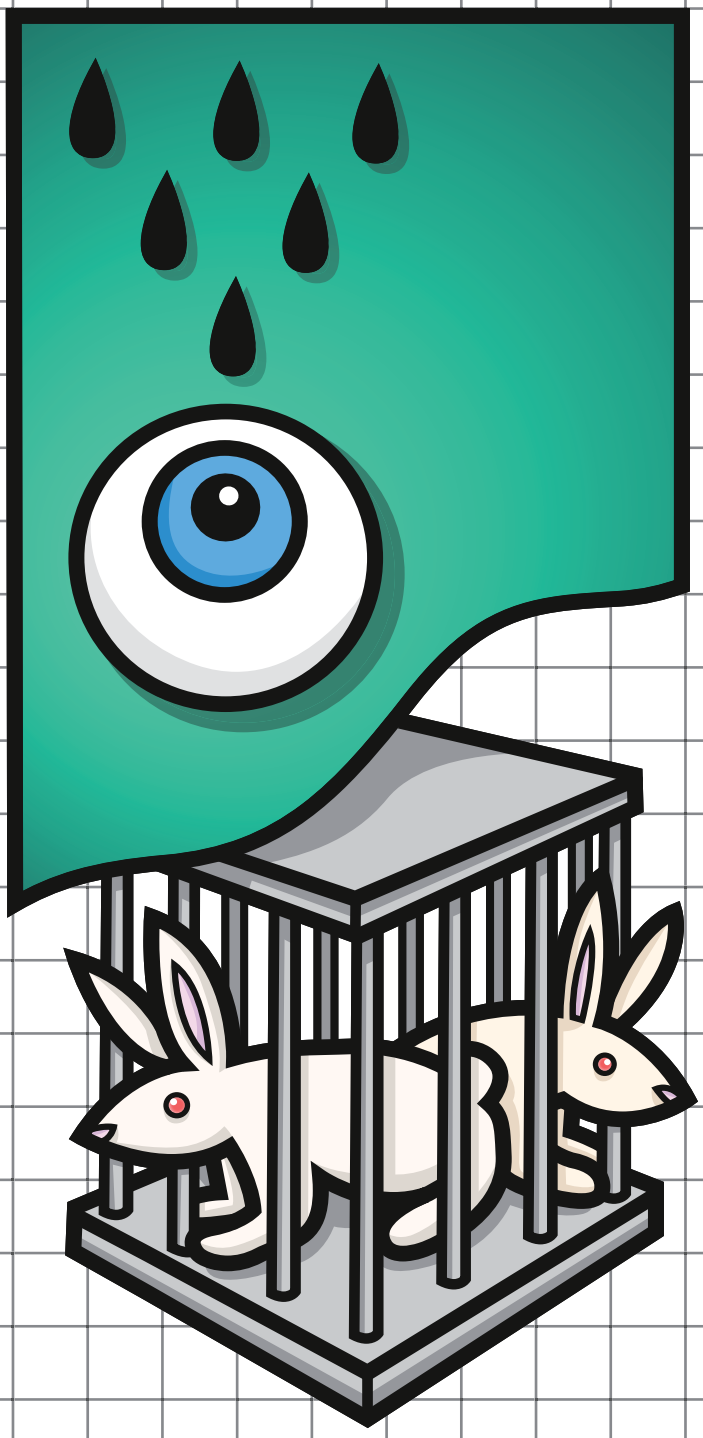
MENK K. PRINSEN



THE ISOLATED CHICKEN EYE TEST TO REPLACE THE DRAIZE TEST IN RABBITS

FROM DEVELOPMENT TO IMPLEMENTATION:
"THE LONG AND WINDING ROAD"

THE ISOLATED CHICKEN EYE TEST TO REPLACE THE DRAIZE EYE TEST IN RABBITS

From development to implementation:
"The long and winding road".

Menk K. Prinsen

THESIS COMMITTEE

Promotors
Prof. Dr R.A. Woutersen
Professor Translational Toxicology
Wageningen University/TNO, Zeist

Prof. Dr C.F.M.Hendriksen
Professor Alternatives to Animal Testing
Utrecht University/Institute for Translational Vaccinology, Bilthoven

Co-promotor
Prof. Dr C.A.M. Krul
Professor Innovative Testing in Life Sciences & Chemistry
Utrecht University of Applied Sciences/TNO, Zeist

Other members
Dr A. van der Lelij, Utrechts Medisch Centrum Utrecht
Prof. Dr R.F. Witkamp, Wageningen University
Prof. Dr B.J. Blaauboer, Utrecht University
Prof. Dr V. Rogiers, Free University Brussel, Belgium

This research was conducted under the auspices of the Graduate School VLAG
(Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

THE ISOLATED CHICKEN EYE TEST TO REPLACE THE DRAIZE EYE TEST IN RABBITS

From development to implementation:
"The long and winding road".

Menk K. Prinsen

Thesis
submitted in fulfilment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus,
Prof. Dr M.J. Kropff,
in the presence of the Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 3 October 2014
at 11 a.m. in the Aula.

Menk K. Prinsen
The Isolated Chicken Eye test to replace the Draize test in rabbits.
From development to implementation: “The long and winding road”, 184 pages.

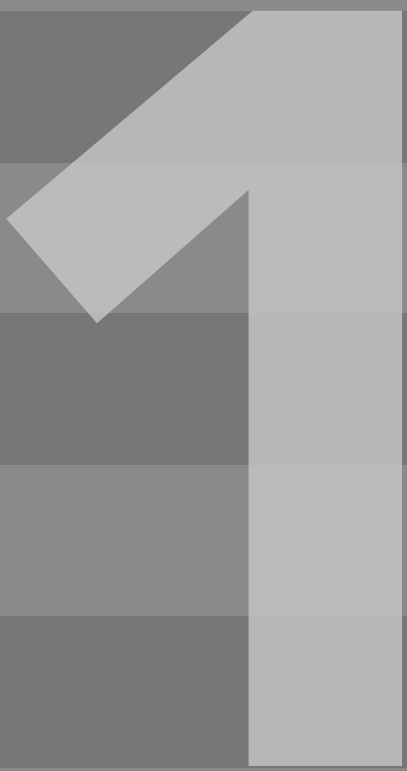
PhD thesis, Wageningen University, Wageningen, NL (2014)
With references, with summaries in Dutch and English
ISBN 978-94-6257-003-0

Voor Cin

Table of contents

Introduction

The Draize Eye Irritation Test and alternatives.



Background

Before industrialization, eye defects were mainly caused by physical trauma or by diseases caused by malnutrition, bacterial infection or parasites. In the twentieth century, when (chemical) industrialization strongly developed, it became apparent that conditions at the workplace could have distinct adverse effects on health and safety of employees. Acute and long-term exposure to a variety of industrial chemicals were responsible for a range of diseases, varying from relatively mild, non-life threatening phenomena, such as dermatitis, to incurable, lethal conditions such as cancer. After World War II, the chemical industry rapidly increased and workers became organized and more concerned with the potential risks they could encounter in the workplace. Consequently, the need for identifying health hazards and worker's protection became an important issue in most industrial countries. Moreover, people could afford more luxury products and the household and personal care industry became more and more innovative using new technologies and (chemical) ingredients. Therefore, an even larger population of people needed to be safeguarded from potential hazardous substances.

To establish the potential risk of exposure of the eyes to compounds, the Food and Drug Administration of the United States (US-FDA) adopted the Draize eye irritation test using rabbits already in 1961.

At first sight, this test is simple and straightforward and provides a useful tool for regulators. However, the controversial character of this type of animal testing became known to the general public – on 15 April 1980, Henry Spira, a Belgian-American advocate, member and founder of the Animal Rights International group bought a full-page advertisement (Figure 1) in the New York Times, with the header: "How many rabbits does Revlon blind for beauty's sake?" – and the need to develop alternative non-animal tests became apparent. Within a year after Spira's advertisment, Revlon had donated $750,000 to a fund to investigate alternatives to animal testing, followed by substantial donations from Avon, Bristol Meyers, Estée Lauder, Max Factor, Chanel, and Mary Kay Cosmetics. These donations led to the creation of the Center for Alternatives to Animal Testing (Wikepedia; Henry Spira).

Figure 1. Spira's advertisement (www.onegreenplanet.org).

The attempts to validate alternative tests for eye irritation in the early nineteen-eighties were considered to be relatively simple by comparing *in vitro* and *in vivo* irritation index scores. What was expected to be a process of several years, however, turned out to be a decades spanning process still not fully completed.

For a large part, this can be attributed to the nature of the *in vivo* test in rabbits, which is more complicated and compromised than originally believed.
This thesis describes the development, performance, validation and acceptance of one of the first alternatives, namely the *in vitro* isolated eye test.

Introduction

The eye

Our eyes are one of the most important sense organs we possess in order to keep in contact with our environment. We have only two eyes and as such we are vulnerable to accidental damage, caused for instance by mechanical trauma or exposure to foreign materials such as chemicals. The exposure to chemicals can be intentional (ophthalmologic formulations and contact lens fluids) or unintentional, at the workplace or at home, using household and/or personal care products.
The natural defence mechanisms against possible damage to our eyes are obvious: protection by i) their embedded position in the eye-socket and protection by the eyelids (blinking, closing of the eyes); ii) lachrymation (production of tears by innervation of the sensory nerves of the cornea) in order to dilute/remove foreign materials; iii) the tear film (protection against bacteria and drying out of the corneal surface), and iv) reflexes (turning away the head, protection of the eyes/head with our hands).

This thesis deals mainly with a critical part of the eye, namely the cornea, the eye's outermost layer and gateway to the perception of light (Figure 2), which is the main target tissue for eye irritation caused by chemical exposure.

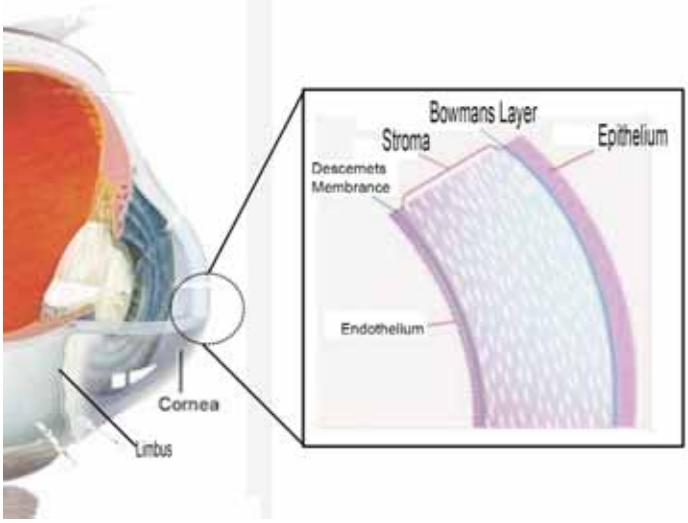


Figure 2. The cornea (www.visionfortomorrow.org).

The composition or histology of the cornea and its functions are described in many handbooks and in general is as follows (NEI, 2011):
“Although the cornea is clear and seems to lack substance, it is actually a highly organized group of cells and proteins. Unlike most tissues in the body, the cornea contains no blood vessels to nourish or protect it against infection. Instead, the cornea receives its nourishment from the tears and aqueous humor that fills the chamber behind it. The cornea must remain transparent to refract light properly, and the presence of even the tiniest blood vessels can interfere with this process. To see well, all layers of the cornea must be free of any cloudy or opaque areas.
The corneal tissue is arranged in five basic layers (Figure 2), each having an important function.
These five layers are:

Epithelium

The epithelium is the cornea's outermost region, comprising about 10 percent of the tissue's thickness. The epithelium functions primarily to: (1) block the passage of foreign material, such as dust, water, and bacteria, into the eye and other layers of the cornea and (2) provide a smooth surface that absorbs oxygen and cell nutrients from tears and distributes these nutrients to the rest of the cornea. The epithelium is filled with thousands of tiny nerve endings that make the cornea extremely sensitive to pain when rubbed or scratched. Cold receptors are abundant in the cornea, but heat and touch receptors are lacking. The part of the epithelium that serves as the foundation on which the epithelial cells anchor and organize themselves is called the basement membrane.

Bowman's Layer

Lying directly below the basement membrane of the epithelium is a transparent sheet of tissue known as Bowman's layer. It is composed of strong layered protein fibers called collagen. Once injured, Bowman's layer can form a scar as it heals. If these scars are large and centrally located, some vision loss can occur.

Stroma

Beneath Bowman's layer is the stroma, which comprises about 90 percent of the cornea's thickness. It consists primarily of water (78 percent) and collagen (16 percent), and does not contain any blood vessels. Collagen gives the cornea its strength, elasticity, and form. The collagen's unique shape, arrangement, and spacing are essential in producing the cornea's light-conducting transparency.

Descemet's Membrane

Beneath the stroma is Descemet's membrane, a thin but strong sheet of tissue that serves as a protective barrier against infection and injuries. Descemet's membrane is composed of collagen fibers (different from those of the stroma) and is made by the endothelial cells that lie below it. Descemet's membrane can regenerate readily after injury.

Endothelium

The endothelium is the extremely thin, innermost layer of the cornea. Endothelial cells are essential in keeping the cornea clear. Normally, fluid leaks slowly from inside the eye into the middle corneal layer (stroma). The endothelium's primary task is to pump this excess fluid out of the stroma. Without this pumping action, the stroma would swell with water, become hazy, and ultimately opaque. In a healthy eye, a perfect balance is maintained between the fluid moving into the cornea and fluid being pumped out of the cornea. Once endothelium cells are destroyed by disease or trauma, they are lost forever. If too many endothelial cells are destroyed, corneal oedema and blindness ensue, with corneal transplantation as the only available therapy."

The Draize eye Irritation test

On 2 November 1944 a manuscript, entitled "Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes" was received for publication by the Journal of Pharmacology and Experimental Therapy. The authors of this article were John H. Draize, Geoffrey Woodard and Herbert O. Calvery from the Division of Pharmacology, Food and Drug Administration, Federal Security Agency, Washington, D.C., USA. It is more than likely that the authors never expected the kind of impact this publication would have on animal experimentation worldwide. Almost seventy years later the name Draize is still inextricably attached to two of the three most disputed toxicity tests commonly used to determine acute toxicity, i.e. the Draize eye irritation test, the Draize skin irritation test and the LD50 (lethal dose) test. The latter two tests fortunately have already been replaced by *in vitro* tests (skin irritation) or by test methods using much less animals and causing less discomfort (LD50). The Draize eye irritation test was first adopted by the US-FDA as part of the safety evaluation of foods, drugs and cosmetics (US-Federal Register, 1961). At that time already, it was recognized that the subjective grading of ocular reactions posed a considerable problem. In order to standardize the scoring and to provide guidance to the observers, an illustrated guide was issued (FDA, 1964; Figure 3). Internationally, the OECD published their first guideline on eye irritation in 1981, which was subsequently adopted by the European Union (EC, 1984).

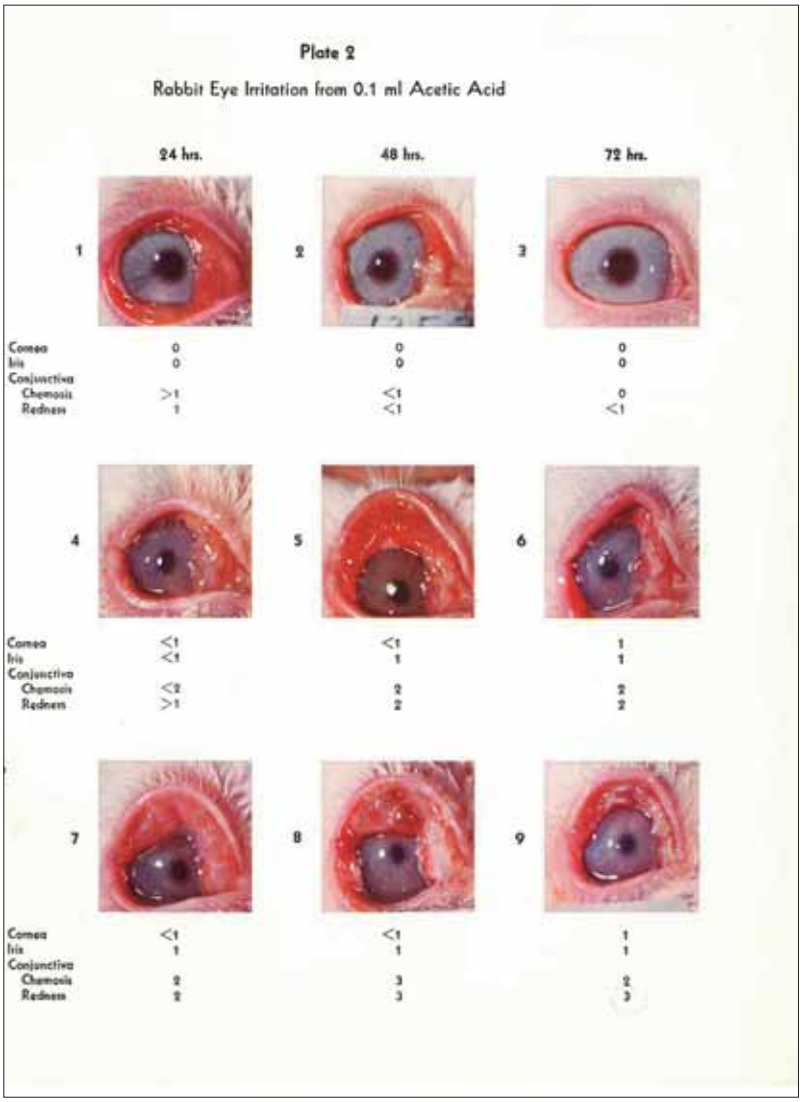


Figure 3. FDA guidance on scoring of ocular lesions; Plate 2 (FDA,1964).

Since then several revisions of the guideline have followed, mostly not affecting the actual exposure procedure, but providing guidance for refinement and reduction of animal use and discomfort (Table 1). Examples are the exemption of testing skin corrosives and substances with pH lower than 2.0 or higher than 11.5, the use of well-validated alternatives as a screen for severe irritancy, and a tiered approach of testing (i.e. starting with one animal and continue only if non-severe irritancy is observed).

Table 1. OECD test guideline no. 405 and its revisions (procedures, interpretation results, ethics and 3 R's).

OECD TG 405	Procedure	Guidance on interpretation of results	Ethical considerations	Three R's
1981	- 0.1 mL or 0.1 g substance - wash out only after 24 hr	- Extrapolation of the results of eye irritation studies in animals to man is valid only to a limited degree. The albino rabbit is more sensitive than man to ocular irritants or corrosives in most cases. - Similar results in tests on other animal species can give more weight to extrapolation from animal studies to man. - Care should be taken in the interpretation of data to exclude irritation resulting from secondary infection.	Local anaesthetics proposed	- Three instead of six rabbits No testing of: - Strongly acidic or alkaline substances - Corrosive or severe skin irritants
1987	- 0.1 mL or 0.1 g substance - wash out only after 24 hr	Identical to 1981 Guidance	Addition of: - Animals showing severe and enduring signs of distress and pain may need to be humanely killed.	Addition of: - severe eye irritants identified in well-validated alternative studies
2002	- 0.1 mL or 0.1 g substance - wash out after 1 hr	Similar to 1981 and 1987 Guidance	Addition of: - End points for humane sacrifice - Tiered testing	Addition of: - Weight-of-the-evidence analysis on the existing relevant data - Conduct of validated and accepted in vitro tests - One rabbit first
2012	- 0.1 mL or 0.1 g substance - wash out after 1 hr	Similar to 1981, 1987 and 2002 Guidance	Addition of: - Extensive directions for the use of topical anaesthetics and systemic analgesics	Addition of: - ICE test (OECD 438) - BCOP test (OECD 437)

The design of the eye irritation test is actually quite simple and straightforward: a rabbit is placed on a worktable and restrained either manually or in a fixation-box (Figure 4). Next, the lower eye-lid is pulled out and the test substance is instilled in the conjunctival cul de sac formed; the upper and lower eye lids are then closed and subsequently held together for at least one second before releasing the animal. The other eye remains untreated and serves as a control.

Figure 4. Instillation of the test substance in the Draize eye test (TNO)

The animal is returned to its cage and is free to remove the material. The control and test eyes are examined (without optical aid) at approximately one hour, and at approximately 24, 48, and 72 hours after treatment. Ocular reactions of the test eye are judged using a scoring scale (Table 2). Residual eye effects are recorded at regular intervals, if necessary up to about 3 weeks after treatment, in order to allow the evaluation of the reversibility or irreversibility of the effects elicited. Liquids are tested in a volume of 0.1 mL and solids (ground to a fine powder) in an amount of 0.1 g or a volume of 0.1 mL. In general, 0.1 mL is the amount the conjunctival cul de sac can hold when the lower eye-lid is pulled out.

Table 2. Draize scheme for grading of ocular lesions in the rabbit.

Tissue	Lesion	Score
Cornea Opacity-degree of density (area most dense taken for reading)	No opacity	0
	Scattered or diffuse areas, details of iris clearly visible	1*
	Easily discernible translucent area, details of iris slightly obscured	2
	Opalescent areas, no details of iris visible, size of pupil barely discernible	3
	Opaque, iris invisible	4
Iris	Normal	0
	Folds above normal, congestion, swelling, circumcorneal injection (any or all of these or combination of any thereof); iris still reacting to light (sluggish reaction is positive)	1*
	No reaction to light, haemorrhage, gross destruction (any or all of these)	2
Conjunctivae - Redness	Vessels normal	0
	Vessels definitely injected above normal	1
	More diffuse, deeper crimson red, individual vessels not easily discernible	2*
	Diffuse beefy red	3
Conjunctivae - Swelling	No swelling	0
	Any swelling above normal (including nictitating membrane)	1
	Swelling with lids about half closed	2*
	Swelling with lids about half closed to completely closed	3

* lowest score considered positive according to US-EPA

Despite the existence of many national guidelines on eye irritation, the exposure procedure and the scoring system for ocular lesions are basically identical. However, the classification systems differ considerably (Tables 3 and 4). In general, four classifications are assigned on the basis of the ocular lesions, viz. not irritating (not classified), mildly irritating, irritating and severely irritating. The EU recognizes three classifications, i.e. not classified, irritating and severely irritating (risk of serious damage to the eye). The existence of these different classification and labelling systems is not favourable for the validation of alternative test methods. Therefore, the implementation of the classification system of the United Nations Globally Harmonized System (UN-GHS; Table 5) in 2007 is considered to be an improvement, although there is still a difference with the system the EU applies.

Table 3. European Union (1993[1]) classification system for eye irritation/corrosion.

Eye effects	R36 (Irritating to eyes)		R41 (Risk of serious damage to eyes)[4]	
	3 animals[2]	6 animals[3]	3 animals[1]	6 animals[2]
Corneal opacity	≥ 2.0, but <3.0	≥ 2.0, but <3.0	≥ 3.0	≥ 3.0
Iris lesion	≥ 1.0, but <2.0	≥ 1.0, but ≤1.5	≥ 2.0	> 1.5
Conjuntiva redness	≥ 2.5	≥ 2.5		
Conjunctiva chemosis	≥ 2.0	≥ 2.0		

[1] Official Journal of the European Communities, L 110 A, Volume 36, 4 May 1993
[2] The classification is assigned if the mean tissue effect (averaged over the 24h, 48h and 72 h time points) exceeds the threshold value in at least two of the three animals.
[3] The classification is assigned if the mean tissue effect (averaged over the three time points and over the six animals) exceeds the threshold value.
[4] R41 is also assigned if, in at least one animal, one of the eye effects has not reversed at the end of the observation period.

Table 4. US-EPA (1998[1]) classification system for eye irritation.

Toxicity categories	Category I	Category II	Category III	Category IV
Eye effects	Corrosive (irreversible destruction of ocular tissue) or corneal involvement or irritation persisting for more than 21 days	Corneal involvement or irritation clearing in 8-21 days	Corneal involvement or irritation clearing in 7 days or less	Minimal effects clearing in less than 24 hours

[1] Health Effects Test Guideline OPPTS870.1000, EPA 712-C-98-189, August 1998

The UN-GHS system subdivides the irritating category (Category 2) into mild irritant (Category 2B) and irritant (Category 2A), whereas the EU only uses the category irritant (Category 2; EC, 2008).

Table 5. UN-GHS (2007[1]) classification system for eye irritation/corrosion.

Eye effects	Category 2A[2]	Category 1[3]
Corneal opacity	≥ 1.0	≥ 3.0
Iris lesion	≥ 1.0	> 1.5
Conjunctiva redness	≥ 2.0	
Conjunctiva chemosis	≥ 2.0	

[1] Globally Harmonised System of Classification and Labelling of Chemicals (UN-GHS). UN, New York and Geneva, 2007
[2] All effects have to be reversible within 21 days of treatment. Subcategory of 2B: mildly irritating to the eyes, i.e. eye effects reversible within 7 days of treatment.
[3] Category 1 is also applicable if, in at least one animal, an eye effect has not reversed, or is not expected to reverse, within 21 days of treatment.

Awareness of alternatives for animal testing

The publication of Russell and Burch in 1959 entitled: "The principles of humane experimental technique" stood at the basis of most initiatives relating to the use and development of alternatives for animal experiments. In their publication they postulated the famous and often cited three R's: Reduction, Refinement and Replacement of animal experiments. Nowadays, the 3 R's have become a mantra for scientists and regulators in research areas involving animal experimentation. The initiatives concerning the Draize eye test mainly involved reduction of the number of animals from six to three per test and replacement by the implementation of non-animal alternatives. Certain aspects of the Draize eye test causing considerable pain and discomfort to the animal were dealt with only at a much later stage, i.e. reduction of the time for a wash-out of the test substance from 24 hours to 1 hour after instillation in 2002, and the use of systemic pain relief and topical sedation in 2012 updates of the OECD guideline 405 (Table 1).

In the early nineteen-eighties, some toxicologists within the TNO-CIVO Toxicology and Nutrition Institute in Zeist, the Netherlands, had growing concerns about the use of experimental animals in toxicity testing. One of them, Drs. H.B.W.M. Koëter, explored the possibilities of introducing alternative test methods for standard acute toxicity tests, such as the Draize eye and skin irritation tests. At that time the Netherlands Society of Toxicology (NVT) started a working group named "Kritische Evaluatie Toxiciteitstesten" (KET; Critical Evaluation of Toxicity Testing) of which Drs. Koëter was a member. In Europe, the European Research Group for Alternatives in Toxicity Testing (ERGATT) was founded to stimulate innovative toxicological research and to act as a counterpart to the John Hopkins Center of Alternatives to Animal Testing in the USA which was founded in 1981.

Several alternatives had already been published varying from simple cell toxicity (cytotoxicity tests), through sperm motility, to damage to the chorioallantois membrane of hen's eggs (HET-CAM; Figure 5).

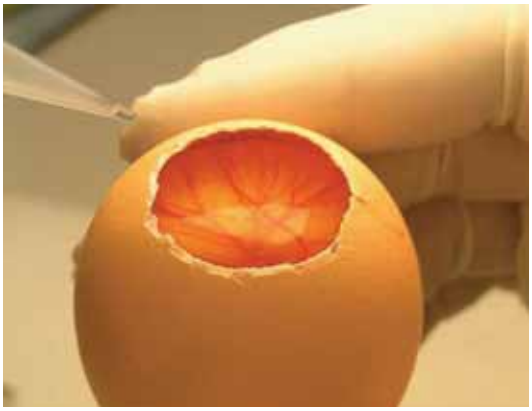


Figure 5. Chorioallantois membrane of the hen's egg (www.eurochemricerche.it).

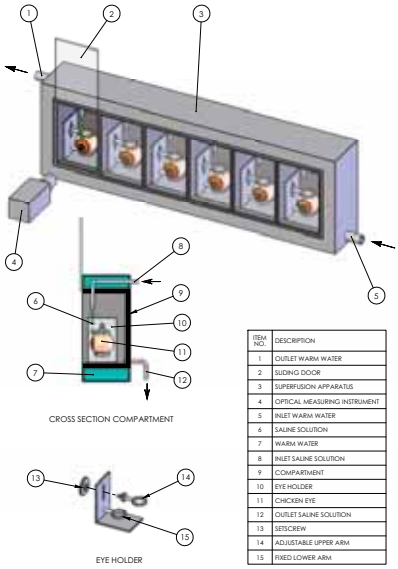
In 1982, TNO-CIVO was invited by the Commission of the European Community to write a report on the reduction of numbers of animals in toxicity testing (Koëter and van Vliet, 1983).Part of the assignment was to make an inventory of alternative test methods used in toxicity testing, and it became apparent that numerous *in vitro* tests had already been developed. Several of these alternative test methods appeared very promising, but standardization and validation had almost never received sufficient attention, because they were developed within a university or company setting, and applied in most cases on a limited scale and not for regulatory purposes.
On the basis of this report, TNO decided to include alternative approaches in the Institute's toxicological research program, which was an important decision at that time. For eye irritation, the policy was not to develop yet another test method but to select one of the most promising existing alternative methods and to focus on further development, standardization and validation in order to develop a method that would be acceptable for regulatory purposes. In addition, recommendations for a tiered approach to eye irritation testing were made, viz. testing skin irritation first, and starting the eye irritation test with one rabbit.
Because the cornea is such a highly relevant target tissue in eye irritation, it was taken as the basic principle for the development of a relevant and practical *in vitro* alternative to the animal test that had been in use as the sole test for the screening of eye irritation worldwide since the early forties of the twentieth century.

Isolated Eye Test method (Rabbit)

In 1981, A.B.G. Burton from Unilever published a method using isolated rabbit eyes for the *in vitro* assessment of severe eye irritants (Burton, 1981). Previously, he had discovered that the measurement of corneal thickness (swelling) by slit-lamp examination provided an objective assessment of eye irritation in the *in vivo* rabbit eye irritation test (Burton, 1972). He had examined 100 different cosmetic formulations in 600 rabbits and found not only a close relationship between the total corneal Draize score and the recorded corneal swelling, but also a relationship between corneal swelling and the conjunctival effects scored subjectively. Around that time another article on the usefulness of slit-lamp examination in the rabbit eye irritation test, including corneal thickness, was published (McDonald, 1973).
Between 1972 and 1981, Burton did not publish further on this subject, but it is assumed that he played with the idea of replacing the live rabbit by isolated rabbit eyes only. In his 1981 publication no further considerations for using isolated eyes were given, but a possible clue may be found in the literature reference he used for the design of the superfusion apparatus (used for maintaining the isolated eyes in good condition), which he had modified from the one described by Mishima and Kudo in 1967. Remarkably, Burton had already referred to publications by Mishima in his 1972 article, and surely have thought about the possibility of using isolated rabbit eyes in a superfusion apparatus at that time. It remains unclear why he did not pursue the use of isolated rabbit eyes sooner.

The idea of using isolated rabbit eyes was very appealing from a scientific point of view. After all one uses an *ex vivo* eye for an eye *in vivo* and, moreover, the parameters measured (corneal swelling, corneal opacity and epithelial cytoxicity by fluorescein dye) are directly comparable to the parameters measured *in vivo* (both in rabbit and in man). Therefore, Koëter proposed to introduce an *in vitro* eye irritation test (with isolated rabbit eyes) as a possible contribution to the reduction of experimental animal use (Koëter and Prinsen, 1984). The project was approved and partly funded by the "Dutch Society for the Protection of Animals" and the foundation "Beauty without Cruelty". Equipment for the Rabbit Enucleated Eye Test (REET; initial name for the Isolated Rabbit Eye Test) was purchased (slit-lamp microscope; Figure 6) or built by the Technical Service Department of TNO-CIVO (superfusion apparatus and eye-holders; Figure 7).

Figure 6. Haag-Streit slit lamp microscope (www.medwow.com).



ITEM NO.	DESCRIPTION
1	OUTLET WARM WATER
2	SLIDING DOOR
3	SUPERFUSION APPARATUS
4	OPTICAL MEASURING INSTRUMENT
5	INLET WARM WATER
6	SALINE SOLUTION
7	WARM WATER
8	INLET SALINE SOLUTION
9	COMPARTMENT
10	EYE HOLDER
11	CHICKEN EYE
12	OUTLET SALINE SOLUTION
13	SETSCREW
14	ADJUSTABLE UPPER ARM
15	FIXED LOWER ARM

Figure 7. Schematic presentation of the superfusion apparatus and eye holder (TNO).

The test method was evaluated by investigating the effects of several substances from the Burton publication (1981). The test method was further validated with 34 substances that had been investigated in the *in vivo* eye irritation test in rabbits as part of the standard toxicity testing at TNO-CIVO in 1983-1984 (Koëter and Prinsen, 1985). During the same period, several other investigators explored the use of isolated rabbit eyes as an alternative for the *in vivo* test. Unilever continued their work (York, 1982), while Shell Research Centre, Sittingbourne, UK, started their research program (Price, 1985). The Institute for Hygiene and Epidemiology (IHE), Brussels, Belgium started their investigations in 1988 (Jacobs, 1988 and 1990).
The research and publications on isolated rabbit eyes resulted in the participation of this test method in the first EC Collaborative Study on the Evaluation of Alternative Methods to the Eye Irritation Test (EC, 1991). In this study, five *in vitro* cell toxicity tests, the REET and the HET-CAM were selected to undergo validation by testing 21 chemicals of different classes in at least 3 different laboratories. Some of the conclusions were: i) The Isolated Rabbit Eye test did not misclassify many non-irritants and also had the capability to discriminate between moderate and severe effects, although irritating (R36) chemicals were underrepresented. ii) The REET produced results which were consistent across all three laboratories and generally correctly predicted the *in vivo* grade. The protocol and the method for calculating final irritancy grades (in validation studies later on called "Prediction Model") needed harmonization before a wider interlaboratory study could be conducted. iii) The test is nearest to the human situation and has the advantage that all types of chemicals can be investigated without the need for testing dilutions, therefore it is easier to interpret than the other assays in this trial. The trial was, however, considered to be too limited to make firm conclusions and it was recommended to perform further interlaboratory trials with a larger number of laboratories and chemicals and according to the principles of GLP. An important remark in the EC report was the fact that toxicological profiles of the chemicals investigated were prepared by collecting and critically evaluating the literature data available, because it was not possible to repeat *in vivo* eye determinations for animal protection considerations. The availability, evaluation and appraisal of *in vivo* eye irritation data and the test method itself constitute the main cause of the exceptionally long, and not yet completed, acceptance of alternative methods to the Draize eye test. Until recently, the *in vivo* data were taken as the "Golden Standard", which in practice meant that the *in vitro* data had to almost exactly match the *in vivo* result.

Isolated Eye Test method (Chicken)

During the introduction and validation of the REET, it was recognized that the use of laboratory rabbit eyes - although available from rabbits used for standard eye and skin irritation tests *in vivo* - was not ideal, especially for laboratories not using rabbits for their experiments. Moreover, the alternative test would still be associated with the use of the laboratory rabbit and eventually, if an alternative would replace the *in vivo* skin and/or eye irritation test, the rabbit as a source for eyes would dry up. The suggestion

had been made that an alternative approach could be to use eyes obtained from rabbit-abattoirs, but use of bovine or chicken eyes was also suggested (Koëter and Prinsen, 1987). Therefore, in 1990, a proposal to examine the suitability of eyes of slaughter-house animals as an alternative for rabbit eyes in the enucleated eye test was submitted to the "Platform Alternatives for Animal Experimentation" (PAD; Platform Alternatieven voor Dierproeven), which was granted after being reviewed by the NWO (Dutch organization for scientific research). During the period October - December 1990, pig and cow slaughterhouses (Hilversum) and chicken slaughterhouses (Breukelen and Nijkerkerveen), all within a 1-hour drive from the test facilities of TNO, were visited to make preliminary investigations concerning the practical aspects of obtaining eyes. In **Chapter 3** the selection of the most suitable eye donor and the development of the method with the selected eye donor species is described. An important aspect was the validation of the test method, i.e. to put it simply: comparing the *in vitro* data with data from the *in vivo* test. Most alternatives are validated with *in vivo* data obtained from literature, a process with many drawbacks, which will be discussed in more detail in **Chapters 7** and **8** of this thesis. The use and suitability of eyes of slaughterhouse animals was first established by testing the same reference chemicals (Prinsen and Koëter, 1993) that had been tested in the EU Collaborative Study on the Evaluation of Alternative Methods to the Eye Irritation Test (EC, 1991). Although the *in vivo* data were obtained from literature this study was considered quite valuable because the *in vitro* data obtained with the slaughter eyes could be directly compared with the *in vitro* data obtained in the REET. Ideally, the *in vitro* test should be performed in parallel with the *in vivo* test, hence enabling a more direct comparison of the data. Fortunately, TNO is also a Contract Research Organization (CRO) and many different substances from various international and national industries were submitted for acute toxicity testing including the eye irritation test. This offered the unique possibility to first test the substances in the isolated eye test prior to the conduct of the *in vivo* eye irritation test.

Other alternatives

In the early nineteen-nineties another alternative method using corneas was developed, namely the Bovine Corneal Opacity test (BCOP; Gautheron, 1992). Gautheron, who worked for Merck, Sharpe and Dhome located in the Auvergne, France used bovine corneas, not *in situ*, but excised from the eye-ball and clamped inside a chamber (Figure 8). At first sight the method appears quite similar to the Isolated Chicken Eye (ICE) test, i.e. using corneas and measuring opacity and fluorescein penetration, but the differences are remarkable. Corneal opacity is measured quantitatively as the amount of light transmission through the cornea. Permeability is measured quantitatively as the amount of sodium fluorescein dye that passes across the full thickness of the cornea, as detected in the medium in the posterior chamber (OECD TG437, 2013). An empirically-derived formula is used to calculate an *In Vitro* Irritancy Score (IVIS = mean opacity value + (15 x mean permeability OD490 value)).

Figure 8. BCOP test chambers with bovine cornea (www.iivs.com)

The BCOP, the ICE test and 7 other test systems were considered to be the most promising alternatives to be further validated and in 1992 the British Home Office (HO) and the Directorate General XI of the European Commission (EC) commissioned a validation study on alternatives to the Draize eye irritation test, to be known as the EC/HO validation study. The first priority was given to evaluate the possibility of identification of substances severely irritating to the eye, while also evaluating the methods for predicting the irritants and non-irritants (Balls, 1995). The methods selected, their principle, expression of results together with the pros and cons are presented in Table 6.

Table 6. Alternative *in vitro* tests for eye irritancy considered most promising (based on EC/HO study).

Alternative	Principle	Expression of results	Pro's	Con's
Red blood cell haemolysis test	Leakage of haemoglobin (H) from red blood cells and denaturation (D)	H_{50} and D values equivalent to MMAS (Modified Maximum Average Score)	- relatively simple set-up - relatively simple performance	- single index score - no direct relation with ocular tissue - no reversibility - testing of non-soluble substances
Neutral red uptake test	Inhibition of neutral red uptake (NRU) into mouse 3T3 cells	NRU_{50} values equivalent to MMAS	- relatively simple set-up - relatively simple performance	- single index score - no direct relation with ocular tissue - no reversibility - extreme PH, non-soluble substances
Fluorescein leakage test	Fluorescein leakage (FL) by damage to the tight junctions of Madin-Darby canine kidney cells	FL_{20} score equivalent to MMAS	- relatively simple set-up - relatively simple performance	- single index score - no direct relation with ocular tissue - no reversibility - viscous materials, extreme PH, non-soluble substances
EYTEX method	Turbidity of reagent	EYTEX Draize equivalent (EDE) score equivalent to MMAS	- relatively simple set-up - relatively simple performance	- single index score - no direct relation with ocular tissue - no reversibility - testing of solids, coloured samples, surfactants, water-solubles - interference/inhibition with matrix
HET-CAM method	Haemorrhage, lysis and coagulation in the chorioallantoic membrane of embryonated chicken eggs	Reaction time for occurrence of haemorrhages, lysis and coagulation within 5 minutes combined into a Q score equivalent to MMAS	- relatively simple set-up - relatively simple performance	- single index score - no direct relation with ocular tissue - limited testing of solids and sticky materials - use of live embryo - subjective scoring - no reversibility
Silicon microphysiometer test	Reduction in the metabolic acidification rate of L929 fibroblasts	MRD_5) values equivalent to MMAS	- assesses functional cell changes	- single index score - no direct relation with ocular tissue - very limited testing of substances (37-48%) - laborious - complex and expensive system
Bovine corneal opacity/ permeability test	Changes in opacity and in permeability of isolated bovine corneas	In vitro irritancy score (IVIS) equivalent to MMAS	- highly standardized - "human" parameters - ocular tissue - eyes relatively easy attainable - objective scoring	- single index score - no direct observation (black box) - cornea excised - thick cornea compared to rabbit/human - laborious - no reversibility - no conjunctival damage - testing of solids, coloured substances
Isolated chicken eye test Isolated rabbit eye test	Corneal swelling, corneal opacity and fluorescein staining of damaged epithelial cells of the cornea	Degree of severity (categories) for each endpoint separately and combination of the three categories into regulatory classification	- eyes easy attainable - relatively simple set-up - relatively simple performance - ocular tissue - "human" parameters - slit-lamp microscopical assessment - objective scoring corneal swelling - direct translation to human ocular damage - all substances can be assayed neat	- no reversibility - no conjunctival damage - subjective scoring opacity, fluorescein retention - experienced observer
Draize rabbit eye test	Corneal opacity, iritis and conjunctival damage of one eye treated in the conjunctival sac	Degree of severity for each endpoint separately and classification on the basis of the most affected tissue (degree and/or persistency)	- simple set-up - simple performance - rabbits easily attainable - large eye - in vivo response including recovery	- unrealistic exposure area (inside eye-lid) - undefined exposure time (seconds to 24 hr) - no conjunctival damage - subjective scoring - experienced observer - animal behaviour influencing eye effects - unrealistic assessment of recovery (no aftercare)

Other alternatives using reconstructed (human) corneal tissue, the so-called 3D models, such as the SkinEthic Human Corneal Epithelium test and the EpiOcular™ test were developed in the late 20th early 21st century and were also validated in several studies. A drawback of these corneal models is that only the epithelial layer of the cornea is reconstituted which might pose a problem in discriminating irritating from moderately/ severely irritating substances.

At present, only the ICE and the BCOP tests are officially adopted by the OECD for Identifying i) Chemicals Inducing Serious Eye Damage (OECD TG 437 and TG 438, 2009) and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage (OECD TG 437 and TG 438, 2013). The Fluorescein Leakage test has been adopted by the OECD for Identifying Ocular Corrosives and Severe Irritants (OECD TG 460, 2012), but with specific limitations, i.e. only applicable to water soluble chemicals

and excluding strong acids and bases, cell fixatives and highly volatile chemicals.

Outline of the thesis

Chapter 2 describes the results of the first validation of the rabbit enucleated eye test (REET) at TNO-CIVO. Substances, already tested in the *in vivo* eye irritation test at the request of various industries and representing the average supply of substances in contract research, were tested in the REET. The test results were used to further develop and optimize the test method and to establish the prediction model for classifying the substances according to their eye irritating properties. **Chapter 3** describes the search for suitable animal species from slaughter-houses as a source for eyes to be completely independent from laboratory rabbits. The most promising candidate, the chicken eye, was further tested with 21 reference compounds to prove its reliability. Thereafter, the isolated chicken eye (ICE) test was immediately incorporated as a prescreen in the routine *in vivo* assessment of eye irritation testing in the frame of contract research at TNO. In **Chapter 4** the successful implementation of the ICE test is described by presenting the parallel *in vitro* and *in vivo* data of a number of substances which represent the average supply of substances to be investigated by a CRO. The ICE test was also used as a stand-alone test, especially by the household and personal care industry which increasingly adopted non-animal testing strategies. The Procter & Gamble Company was one of these companies that employed the ICE test to their eye irritation safety program, and **Chapter 5** describes the application of the ICE test to their domain of household cleaning products. **Chapter 6** deals with investigations in the search for additional parameters that could be helpful to discriminate between the different irritancy levels in the ICE. Histopathology of the cornea with different staining techniques and the determination of the corneal "Depth-of-Injury" could provide more decisive data, especially in those borderline cases between irritant and severely irritant. The need for accepted alternative methods led to international validation studies involving several promising alternatives. In **Chapter 7** the use of the Modified Maximum Average Score (MMAS) as the sole parameter for evaluation of *in vivo* eye irritation is discussed. One of the most comprehensive international validation studies with nine alternative methods including the ICE was held in the mid nineteen-nineties. The results, however, were very disappointing and one of the reasons for that was believed to be the use of the MMAS. Recommendations for handling of data in future validation studies are given.

Obtaining regulatory acceptance of *in vitro* methods for eye irritation has been, and still is, a time-consuming activity. The main obstacle is that regulatory bodies such as ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods) demand that any alternative method must have an almost perfect match with the *in vivo* eye irritation test. **Chapter 8** discusses the problems that developers of alternative methods for eye irritation are facing when they are urged to strictly use the *in vivo* eye irritation data to validate their method and to gain regulatory acceptance. **Chapter 9** contains the general discussion of the results and conclusions on the

development, validation and practical application of the ICE test with the emphasis on the reasons for the long and still continuing process of regulatory acceptance.

References

Balls, M., Botham, P.A., Bruner, L.H., Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology in Vitro* 9, 871-929.

Burton A.B.G. (1972). A method for the objective assessment of eye irritation. Food and *Cosmetics Toxicology* 10, 209-217.

Burton A.B.G., York M., Lawrence, R.S. (1981).The *in vitro* assessment of severe eye irritants.; *Food and Cosmetics Toxicology* 19, 471-480.

Draize, J.H., Woodard, G., Calvery, H.O. (1944). Methods for study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *Journal of Pharmacology and Experimental Therapeut*ics 82, 377-390.

EC (1984). *Directive 67/548 (6th adaption)*; Annex V, Part B: Methods for the Determination of Toxicity B.5. Acute toxicity, eye irritation. Official Journal of the European Community 27, L251, 109.

EC (1991). Collaborative study on the evaluation of alternative methods to the eye irritation test. *Document XI/632/91-V/E/1/131/91*, part I (pp. 54) and part II (pp. 196). Brussels: European Commission.

Gautheron, P., Dukic, M., Alix, D., Sina, J.F. (1992). Bovine corneal opacity and permeability test: An *in vitro* assay of ocular irritancy. *Fundamental and Applied Toxicology* 18, 442-449.

Jacobs, G., Martens, M. (1988). The enucleated eye test: a comparison of the use of ultrasonicand optic pachometers. *Toxicology In Vitro* 2(4), 253-256.

Jacobs, G., Martens, M. (1990). Quantification of eye irritation based upon *in vitro* changes of corneal thickness. *ATLA* 17, 255-262.

Koëter, H.B.W.M., van Vliet, J.C.J. (1983). Possibilities for reduction of the use of laboratory animals in toxicity testing. *TNO report no.* V83.249/000231, September 1983, Zeist, the Netherlands.

Koëter, H.B.W.M., Prinsen, M.K. (1984). Introduction of an *in vitro* eye irritation test as a possible contribution to the reduction of the number of animals in toxicity testing. *TNO report no.* V84.150/140322, May 1984, Zeist, the Netherlands.

Koëter, H.B.W.M., Prinsen, M.K. (1985). Comparison of *in vivo* and *in vitro* eye irritation test systems: A study with 34 substances. *Alternative Methods in Toxicology* 3, Chapter A9. Mary Ann Liebert, Inc., publishers, New York.

Koëter, H.B.W.M., Prinsen, M.K. (1987). Validation of an *in vitro* eye irritation study; A first step. *Alternative Methods in Toxicology* 5, Chapter E5. Mary Ann Liebert, Inc., publishers, New York.

McDonald, T.O., Baldwin H.A., Beasley, C.H. (1973). Slit-lamp examination of experimental animal eyes. *Journal of the Society of Cosmetic Chemists* 24, 163-180.

Mishima, S., Kudo, T. (1967). *in vitro* incubation of rabbit cornea. *Investigative Ophthalmology* 6 (4), 329-339.

National Eye Institute (NEI) (2011). Facts about the cornea and corneal disease, available at http://www.nei.nih.gov/health/cornealdisease (accessed April 2011).

OECD (1981). *OECD Guideline for Testing of Chemicals* No. 405: Acute eye irritation/corrosion, adopted 12 May 1981. Organisation for Economic Co-operation and Development, Paris.

OECD (2009). *OECD Guideline for Testing of Chemicals* No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying Ocular Corrosives and Severe Irritants, adopted 7 September 2009. Organisation for Economic Co-operation and Development, Paris.

OECD (2009). *OECD Guideline for Testing of Chemicals* No. 438: Isolated Chicken Eye Test Method for Identifying Ocular Corrosives and Severe Irritants, adopted 7 September 2009. Organisation for Economic Co-operation and Development, Paris.

OECD (2012). *OECD Guideline for Testing of Chemicals* No. 405: Acute eye irritation/corrosion, adopted 2 October 2012. Organisation for Economic Co-operation and Development, Paris.

OECD (2012). *OECD Guideline for Testing of Chemicals* No. 460: Fluorescein Leakage Test Method for Identifying Ocular Corrosives and Severe Irritants, adopted 2 October 2012. Organisation for Economic Co-operation and Development, Paris.

OECD (2013). *OECD Guideline for Testing of Chemicals* 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, adopted 26 July 2013. Organisation for Economic Co-operation and Development, Paris.

OECD (2013). *OECD Guideline for Testing of Chemicals* 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, adopted 26 July 2013. Organisation for Economic Co-operation and Development, Paris.

Price, J.B., Andrews, I.J. (1985). The *in vitro* assessment of eye irritation using isolated eyes. *Food and Cosmetic Toxicology* 23(2), 313-315.

Prinsen, M.K., Koëter, H.B.W.M. (1993). Justification of the Enucleated Eye Test with eyes of slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits. *Food and Chemical Toxicology* 31(1), 69-76.

Russell W.M.S., Burch R.L. (1959). *The Principles of Humane Experimental Technique*. South Mimms, Potters Bar, Herts:Universities Federation for Animal Welfare.

United Nations-Economic Commission for Europe (UN/ECE) (2007). Globally Harmonised System of Classification and Labelling of Chemicals (GHS). UN, New York and Geneva, 2007. Available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html.

US-Federal Register (1961). *Title 21 Food and Drugs*, Part 191 Hazardous substances, § 191.12 Test for eye irritants.

US-FDA (1964). Illustrated guide for grading eye irritation by hazardous substances. Washington, D.C. 20204, USA.

York, M., Lawrence, R.S., Gibson, G.B. (1982). An *in vitro* test for the assessment of eye irritancy in consumer products – preliminary findings. *International Journal of Cosmetic Science* 4, 223-234.

Implementation and validation of the Rabbit Enucleated Eye Test

Herman B.W.M. Koëter and Menk K. Prinsen.
Comparison of *In Vivo* and *In Vitro* Eye Irritancy Test Systems: A Study With 34 Substances.

Alternative Methods in Toxicology Volume 3, In Vitro Toxicology: A Progress Report From The Johns Hopkins Center for Alternatives to Animal Testing. Editor Alan M. Goldberg. Chapter A9. Mary Ann Liebert, Inc., New York, 1985

2

Comparison of *In Vivo* and *In Vitro* Eye Irritancy Test Systems: A Study With 34 Substances

HERMAN B.W.M. KOËTER and
MENK K. PRINSEN

Department of Biological Toxicology
Institute CIVO-Toxicology and Nutrition TNO
Zeist, The Netherlands

IN VIVO AND *IN VITRO* EYE IRRITANCY TEST SYSTEMS

ABSTRACT

34 Different substances were tested in both the in vivo eye irritancy test in rabbits and an in vitro method based on the use of isolated rabbit eyes. Comparison of the results revealed that
a. 82% of all substances showed similar effects in both test systems,
b. similar effects were observed not only in the category of severe irritants but occurred at all levels of irritancy,
c. the isolated eye test was more sensitive than the in vivo test for 12% of the substances; these substances were all negative in the in vivo test but appeared to be skin irritants in a primary skin irritation test,
d. for only 6% (2 substances) the isolated eye test was less sensitive than the in vivo test.
It was concluded that the isolated eye test is a sensitive and useful test system. It should therefore seriously be considered as a valid screen for the testing of eye irritating potential of chemicals for which registration is required.

INTRODUCTION

Information on possible adverse effects of substances on the eye is considered essential for the safety evaluation and classification of chemicals and of toiletries, cosmetics and household products. Therefore, in vivo eye irritation tests as described by Draize, Woodard and Calvary (1) are still an essential part of the basic requirements for admission or registration of all kinds of chemicals in most "western" countries. This implies that the so-called Draize test, which is vigourously attacked by anti-vivisectionists and animal-protectors and is often used to illustrate the burdening degree of discomfort to which animals may be exposed, is as yet indispensable. Only validated alternative test methods may possibly supersede the traditional test to some degree.

A rather promising in vitro method for the screening of substances for eye irritancy, which is based on the use of isolated rabbit eyes has been described by Burton et al. (2). To obtain information on the sensitivity and reliability of this test sys-

tem, both the in vivo Draize test and the newly introduced in vitro method were conducted with 34 different substances.
This paper describes this comparative study.

ANIMALS

New Zealand White male rabbits were used in both test systems. Since the intention of this project is to reduce the number of animals in eye irritancy testing, only rabbits that had already been used in primary skin irritation- or eye irritation studies were used as eye donors for the in vitro tests. This implies that all observations in the in vitro tests were done in animals that would have been killed for different reasons anyway.
To achieve the best possible comparison of results from the in vivo- with those from the in vitro test, both tests were preferably conducted successively in the same animals or simultaneously in different animals. Only animals that were in good health and free of any eye defects were used.

MATERIAL AND METHODS

Substances

During the period January 1983 - August 1984, a selection was made of materials tested at the request of various sponsors in the in vivo eye irritation test. The substances were selected on the basis of physical characteristics and pH in such a way that ratios for these physico-chemical properties were obtained that are usually experienced in our laboratory. The selected substances represent not only toiletries and cosmetics but also industrial chemicals and household products.

All substances are listed in table 1 as code numbers since studies with these substances were carried out under contract and thus are confidential.

TABLE 1

Physico-Chemical Characteristics of Substances Tested In Vivo and In Vitro

Substance	pH	Physical characteristic	Substance	pH	Physical characteristic
1	7.3	hydrophilic	18	12.2	hydrophilic
2	6.6	hydrophilic	19	1.8	hydrophilic
3	6.5	hydrophilic	20	-	hydrophilic
4	6.5	hydrophilic	21	-	solid
5	-	solid	22	-	solid
6	12.5	hydrophilic	23	-	solid
7	13.5	hydrophilic	24	-	hydrophobic
8	-	hydrophobic	25	-	hydrophobic
9	-	hydrophobic	26	-	solid
10	-	hydrophobic	27	5.0	hydrophilic
11	-	hydrophobic	28	7.2	hydrophilic
12	-	hydrophobic	29	7.2	hydrophilic
13	-	hydrophobic	30	-	solid
14	-	hydrophobic	31	-	solid
15	-	hydrophobic	32	-	solid
16	5.7	hydrophilic	33	-	hydrophobic
17	-	solid	34	6.0	hydrophilic

Experimental Design

All in vivo eye irritation studies were conducted either according to the OECD-guidelines for testing of chemicals (3) or according to the FDA-guidelines for testing hazardous substances (4). For each substance, the in vivo irritancy grade was assessed by applying the classification criteria defined by the FDA (4).

The in vitro test method was essentially similar to that described by Burton et al. (2). The superfusion apparatus was slightly modified by applying somewhat different dimensions for easier handling of the eyes, and by using black perspex sliding doors for each compartment. Further, the water-jacket was improved by enclosing the compartments on three sides instead of two. For holding the enucleated eyes a different clamp was designed consisting of an upper and a lower stainless steel ring-shaped arm which are both mounted on a stand. The clamp was designed to hold one eye in place in such a way that any undesirable pressure could be avoided. Per test 4 eyes were treated and 2 eyes served as controls. Liquid substances were applied to the eyes in amounts of 0.1 ml while solids were tested by dusting 100 mg onto the eyes. Five to ten seconds after application the corneal surface was rinsed thoroughly with approximately 20 ml of isotonic saline. All eyes were examined with a slit-lamp microscope just before treatment and at regular intervals after application of the test substance.

For the assessment of possible effects, the following criteria were applied:

1. Permeability: Permeability was measured by applying a 2% Fluorescein sodium solution to the surface of the cornea for a few seconds followed by rinsing with isotonic saline.

Permeability was measured before treatment and at t = 30 min. The following scoring system was used:

0 = none or a few cells permeable

1 = small number of cells permeable

2 = individual cells and areas of the cornea permeable

3 = entire cornea permeable

The final score for permeability, being the sum of scores for each of the 4 eyes was interpreted as follows:

1-5 : slight effect

6-9 : moderate effect

10-12: severe effect.

2. Corneal opacity: eyes were examined for corneal opacity at t = 30, 75, 120, 180 and 240 min. The following scoring system was used:

0 = no effect or negligible effect

1 = slight degree of corneal opacity

2 = moderate degree of corneal opacity, but details in the iris are still visible

3 = marked degree of corneal opacity: details in the iris are not visible.

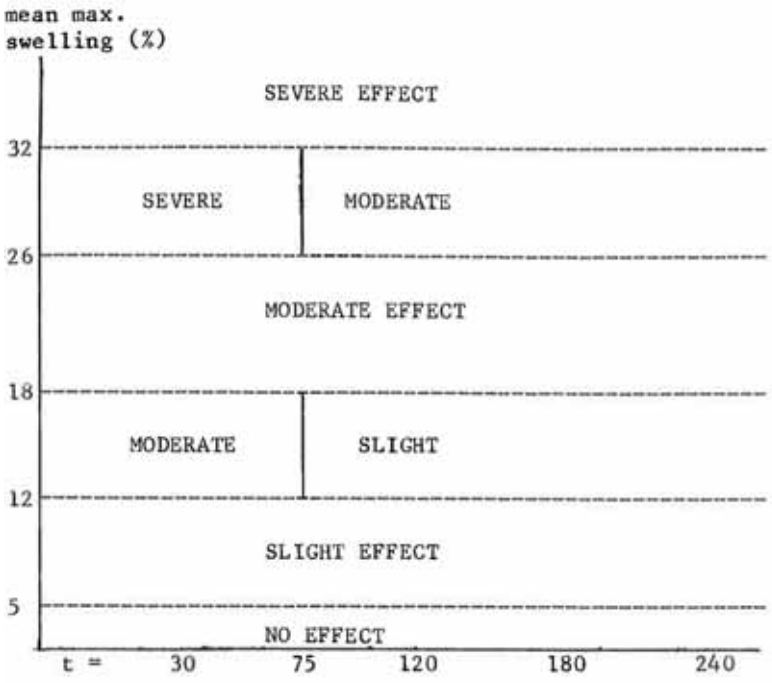
The final score for corneal opacity was similarly calculated and interpreted as described for permeability.

3. Corneal swelling: corneal thickness was measured with a depth-measuring device, mounted on the slit-lamp microscope. Corneal thickness was measured just before treatment and at t = 30, 75, 120, 180 and 240 min. Corneal swelling was expressed as: [(corneal thickness at time t/corneal thickness before treatment)-1] x 100%.

The interpretation of the observed swelling is given in the scheme on the next page and was based not only on the mean maximum swelling for all 4 eyes but also on the time of occurrence.

4. Other effects: these include "pitting" of corneal epithelial cells, "loosening" of epithelium, "roughening" of the corneal surface and "sticking" of the test substance to the cornea. The severity of these findings varied and combinations of effects occurred. The final score for "other effects" was subjective to the interpretation of the investigator and represented the mean value of all 4 eyes.

For each substance, the ultimate in vitro irritancy grade was assessed by averaging the final scores of the 4 criteria.



RESULTS

A comparison of irritancy grades as concluded from in vivo studies with those assessed in the in vitro tests is given in table 2.

From the results it appeared that for 28 substances (82%) similar irritancy grades were assessed in both test systems. It further appeared that similar irritancy grades occurred not only in the category of severe irritants but at all levels of irritancy.

TABLE 2

Comparison of Irritancy Grades as Concluded from In Vitro Scores with those Obtained from the In Vivo Tests

Sub-stance	irritancy grade in vitro	in vivo
1	not irritant	not irritant
2	slight	slight
3	slight	slight
4	moderate/severe	severe
5	slight	slight
6	severe	severe
7	slight	severe
8	severe	severe
9	slight	not irritant
10	negligible	not irritant
11	not irritant	not irritant
12	not irritant	not irritant
13	moderate	moderate
14	slight	not irritant
15	moderate	not irritant
16	not irritant	not irritant
17	severe	severe
18	slight	slight
19	negligible	not irritant
20	not irritant	not irritant
21	slight	slight
22	negligible	not irritant
23	negligible	not irritant
24	negligible	not irritant
25	severe	severe
26	not irritant	not irritant
27	negligible	not irritant
28	not irritant	not irritant
29	not irritant	not irritant
30	slight	severe
31	slight	slight
32	negligible	not irritant
33	moderate	not irritant
34	severe	severe

The in vitro isolated eye test was more sensitive than the in vivo test for 4 substances (12%). Although these substances were all negative in the in vivo eye irritation test, they appeared to

be moderate to severe irritants in a primary skin irritation test. For 2 substances (no. 7 and 30), which is less than 6%, the isolated eye test was less sensitive than the in vivo test.

DISCUSSION

The differences in response of some moderate or severe skin irritants in the in vivo- and the in vitro test again give rise to the discussion whether or not there is a correlation between primary skin irritation and eye irritation. A rather strong positive correlation was reported by Smyth et al. (5), while Williams (6) suggested that any prediction of eye irritancy on the basis of skin irritating properties is misleading. It is also believed that the evaluation of ocular irritation in vivo would be more valuable when objective criteria such as corneal swelling would be included (7).

As it is still not clear whether or not skin irritancy test results are predictive of ocular irritancy potential, we feel that the 4 moderate or severe skin irritants that appeared to be positive in the in vitro test but were negative in the in vivo test should not merely be considered as false positive results.

Only 2 substances revealed lower scores in the in vitro test (slightly irritant) than in the in vivo study. In the in vivo test these substances were considered severe irritants for reasons of persistence of effects.

It was concluded that the in vitro isolated eye test is a sensitive and useful test system. Therefore, the test should seriously be considered as a valid screen for the testing of eye irritating potential of chemicals for which registration is required by national or international regulations such as TSCA (USA) and the 6th Amendment of Directive 67/548 of the EEC. A confirmation of negative results in the in vitro test by also

conducting the in vivo test with a limited number of animals (max. 3) should only be considered when regular eye-contact might be expected.

ACKNOWLEDGMENTS

This project was supported by grants of: Institute CIVO-Toxicology and Nutrition TNO, the Dutch Society for the Protection of Animals, and the Dutch Beauty Without Cruelty Foundation.

REFERENCES

1) Draize, J.H., G. Woodard and H.O. Calvary
J. Pharmacol. Exp. Ther. 82: 377, 1944

2) Burton, A.B.G., M. York and R.S. Lawrence
Fd. Cosmet. Toxicol. 19: 471, 1981

3) OECD-Guidelines for testing of chemicals
Section 4: Health effects, 1981

4) Code of Fed. Regulations
FDA method for testing primary irritant substances
16 CFR 1000.1 paragraph 1500.41, p. 294, 1980

5) Smyth, H.F., Jr., C.P. Carpenter, C.S. Weil, U.C. Pozzani and J.A. Streigel
Am. Ind. Hyg. Ass. J. 23: 95, 1962

6) Williams, S.J.
Fd. Chem. Tox. 22: 157, 1984

7) Conquet, Ph., G. Durand, J. Laillier and B. Plazonnet
Tox. Appl. Pharmacol. 39: 129, 1977

Development of the Chicken Enucleated Eye Test

M.K. Prinsen and H.B.W.M. Koëter.
Justification of the Enucleated Eye Test with eyes of slaughterhouse animals as an alternative to the Draize Eye Irritation Test with rabbits.

Food and Chemical Toxicology Vol. 31, No. 1, pp. 69-76, 1993

3

Fd Chem. Toxic. Vol. 31, No. 1, pp. 69–76, 1993
Printed in Great Britain.

0278-6915/93 $6.00 + 0.00

JUSTIFICATION OF THE ENUCLEATED EYE TEST WITH EYES OF SLAUGHTERHOUSE ANIMALS AS AN ALTERNATIVE TO THE DRAIZE EYE IRRITATION TEST WITH RABBITS

M. K. PRINSEN* and H. B. W. M. KOËTER†
TNO Toxicology and Nutrition Institute, PO Box 360, 3700 AJ Zeist, The Netherlands

(Accepted 9 September 1992)

Abstract—The enucleated eye test (EET) with the isolated eye of rabbits has been recognized as a valuable alternative to the Draize test, because it represents a test system nearest to the *in vivo* test, without the need to use live animals. In this *ex vivo* bioassay, three parameters are measured to detect possible adverse eye effects, namely corneal swelling, corneal opacity and fluorescein retention. The measurement of corneal swelling in this assay guarantees a highly objective and discriminative parameter. In combination with the detailed observation of corneal opacity and fluorescein retention, a reliable evaluation of the eye irritation potential of test materials is achieved. However, laboratory animals are still necessary as eye donors. The use of slaughter animals, such as the cow, the pig and the chicken, as possible eye donors for the EET was therefore examined. From these candidates, the chicken appeared to be the most practicable. 21 reference compounds, ranging from non-irritant to severe irritant, which had been tested previously in a validation study on alternative test methods for eye irritation testing, sponsored by the Commission of the European Communities, were examined in the Chicken Enucleated Eye Test (CEET). When compared with the *in vivo* EC classification, the CEET correctly classified each of the compounds that must be labelled in the EC as irritant (R36) or severely irritant (R41). In addition, since the CEET recognizes three levels of irritancy rather than two (as in the case of the EC classification) a small number of the compounds were recognized as slightly irritant, which according to the EC classification need not be labelled. It was concluded that this *ex vivo* test system is highly accurate in the assessment of eye irritation potential without the use of laboratory animals.

INTRODUCTION

The enucleated eye test (EET) with isolated eyes of rabbits was introduced by Burton *et al.* (1981) as a prescreen for the *in vitro* (*ex vivo*) assessment of severe eye irritants. In this test method, parameters such as corneal thickness, corneal opacity and fluorescein retention are used to disclose possible adverse eye effects of test materials. Earlier, Burton (1971) had shown that measurement of corneal thickness in the rabbit *in vivo* was an objective parameter, which correlated well with the eye lesions observed. In 1983, the EET method was introduced in our laboratories and was validated against the Draize eye test with live animals by testing 34 compounds in both tests (Koëter and Prinsen, 1985). Almost at the same time, the method was introduced at Shell Research Ltd, UK, by Price and Andrews (1985). After testing 60 compounds in the EET, these authors also concluded that the method can be used for a reliable prediction of the potential of chemicals to cause ocular injury. In 1988–1989, an initial collaborative study on the evaluation of alternative methods to the eye irritation test was initiated, sponsored by the Commission of the European Communities (CEC). In this study 10 laboratories participated and 21 reference compounds were examined in the EET, the hen's egg chorioallantoic membrane test, the neutral red uptake test, the agarose overlay test, the total protein content test, the neutral red release test and the lactic dehydrogenase test. From the results (CEC, 1991) it was concluded that the EET provided results that were consistent across the laboratories carrying out this test, and generally predicted the correct *in vivo* grade of the compounds. Furthermore, it was concluded that this method is the nearest to the human situation, and has the advantages that all types of chemicals can be assayed without dilution.

Despite these advantages, the main argument against its acceptance could be that, in those cases where there are no rabbit slaughterhouses to supply the need for human consumption of rabbit meat, laboratory animals are still needed as eye donors. Even under the circumstances where the donor laboratory rabbits would be obtained from other (usually dermal irritation) studies, and would have been killed anyway, the use of additional animals

*To whom correspondence should be addressed.
†Present address: OECD, Environmental Health and Safety Division, 2 rue André Pascal, 75775 Paris Cedex 16, France.

Abbreviations: CEC = Commission of the European Communities; CEET = chicken enucleated eye test; EC = European Community; EET = enucleated eye test; OECD = Organisation for Economic Co-operation and Development; SDS = dodecyl sulphate sodium salt.

cannot fully be excluded. Therefore, the possible use of various, more current species of slaughter animals as possible eye donors for the EET was examined. The suitability and sensitivity of the donor species selected from this exercise were confirmed in a comparative test with the same 21 chemicals used for the EC collaborative study. The comparative test was performed according to the principles of Good Laboratory Practice (OECD, 1991).

MATERIALS AND METHODS

Test materials. The 21 chemicals already used for the EC collaborative study were provided by FRAME (Fund for the Replacement of Animals in Medical Experiments) through Aldrich Chemical Company Limited (Gillingham, UK), and included acetic acid (at least 99% pure), polyoxyethylene (23) lauryl ether (Brij 35, purity not specified), benzalkonium chloride (purity not specified), dimethyl sulphoxide (99.9% pure), sodium fluorescein (water soluble, 70% dye content), glycerol (99% pure), triacetin (99% pure), mercury(II)chloride (99.5% pure), silver(I)nitrate (at least 99% pure), sodium hydroxide (at least 97% pure), toluene (99.9% pure), triethanolamine (99% pure), *n*-hexane (99% pure), chloroform (99.8% pure), 2-methoxy ethanol (anhydrous 99.9% pure), 1-butanol (99% pure), acetaldehyde (99% pure), 2-butoxyethyl acetate (99% pure), dodecylsulphate sodium salt (SDS, 70% pure), dibutyltin dichloride (97% pure) and tributyltin chloride (96% pure). The chemicals were tested undiluted with the exception of acetic acid, silver(I)nitrate, sodium fluorescein, and sodium hydroxide, which were tested as 10, 3, 20 and 1% dilutions in demineralized water, respectively.

Species comparison study. Three species were selected as possible donor candidates for the EET, namely cow, pig and chicken. Preliminary investigations were carried out to determine the suitability of each species as an eye donor. These investigations included inventory of slaughterhouses within a range of 1 hr from the testing facility, availability of the eyes, and dissection, examination and processing of the eyes.

To investigate the availability and suitability of the cow as an eye donor, eyes were collected in a slaughterhouse (Hilversumse Slacht B.V., Hilversum, The Netherlands). Immediately after the animals had been killed by means of a humane cattle-killer and subsequent bleeding, the eyes were dissected by the slaughterhouse staff working on the process line, who had received detailed instructions. The enucleated eyes were each placed in a plastic, saline-filled cup with the cornea positioned downwards; the cups were placed in plastic containers with closed lids and transported to the testing facility for further processing and examinations.

To investigate the availability and suitability of the pig as an eye donor, eyes were collected in two slaughterhouses (Hilversumse Slacht B.V., Hilversum, The Netherlands, and Deen's B.V., Bussum, The Netherlands). The eyes had to be collected shortly after the animals had been killed by means of electroshock and subsequent bleeding, and before they reached the next station on the process line, which is a hot water bath. In contrast with the dissection procedure of the bovine eyes, where it was impracticable for the investigator to intervene in the process line, the eyes could be removed by the investigator, thus ensuring proper dissection. The nictitating membrane was firmly drawn away from the eyeball with surgical forceps and held in this position during the entire dissection procedure. The conjunctivae between the nictitating membrane and the eyeball were then cut with curved scissors. Next, the extraorbital muscles, optic nerve and remainder of the conjunctivae were cut, and the eyeball was removed from the orbit. Dissections were performed with extreme care to avoid touching the surface of the cornea. Cutting the optic nerve too close to the eyeball, which could result in rupture and consequent loss of intraocular pressure, was also avoided. Immediately after dissection, the enucleated eye was placed in a 30-ml glass bottle completely filled with isotonic saline at ambient temperature and transported to the testing facility for further processing and examination.

To investigate the availability and suitability of the chicken as an eye donor, a poultry slaughterhouse relatively close to the laboratory (v.d. Bor, Nijkerkerveen, The Netherlands) was selected. Immediately after sedation by electric shock and incision of the neck for bleeding, and before they reached the next treatment location on the process line, the birds' heads were cut off. The heads were placed in small plastic boxes (three heads per box) on a bed of paper tissues moistened with isotonic saline. During transportation to the testing facility, the heads were kept at ambient temperature. Within 2 hr after death, eyes were carefully dissected and placed in a superfusion apparatus for further treatment and examination.

Comparative study. This study was conducted with chicken eyes only and generally followed the same procedure as that used for the EET with rabbit eyes. Approximately 7-wk-old male or female chickens (Ross Spring Chickens; v.d. Bor, Nijkerkerveen, The Netherlands), body weight range approximately 2.5–3.0 kg, were used as eye donors. For each compound to be tested, the heads of 12 birds were collected. The same procedure for collection of the heads was followed as for the species comparison study. Within 2 hr after death, eyes were carefully dissected and placed in a superfusion apparatus using the following procedure.

First, the eyelids were carefully removed without damaging the cornea and a small drop of fluorescein sodium BP 2% (w/v) (Minims, Smith & Nephew Ltd, Romford, UK) was applied to the corneal surface for

a few seconds and subsequently rinsed off with isotonic saline of ambient temperature. Next, the head with the fluorescein-treated cornea was examined with a slit-lamp microscope (Slit-lamp 900 BM, Haag-Streit AG, Liebefeld-Bern, Switzerland), to ensure absence of any damage. If undamaged, the eye was removed from the orbit very carefully, paying special attention that the optical nerve was not cut too short. The enucleated eye was placed in a stainless-steel clamp with the cornea positioned vertically and then transferred to a chamber of the superfusion apparatus (modified from Burton *et al.*, 1981). The clamp holding the eye was positioned in such a way that the entire cornea was supplied with isotonic saline from a bent, stainless-steel tube, at a rate of approximately 0.10–0.15 ml/min (peristaltic pump, Desaga STA 131900, Heidelberg, Germany). The six chambers of the superfusion apparatus as well as the saline were temperature controlled at 32 ± 1.5°C (water pump, Thermomix 1441, B. Braun Melsungen AG, Melsungen, Germany).

After six eyes had been selected and placed in the superfusion apparatus, these eyes were examined again with the slit-lamp microscope to ensure that they were not damaged during the procedure. Corneal thickness was measured using the Depth Measuring Attachment no. II for the Haag-Streit slit-lamp microscope. Thickness of the cornea was expressed in instrument units. An accurate measurement was taken at the corneal apex of each eye. Eyes with a corneal thickness deviating by more than 10% of the average corneal thickness of the six eyes, or eyes that were unacceptably stained with fluorescein, indicating that the cornea was permeable, or eyes that showed any other signs of damage, were rejected as test eyes and were replaced. After an equilibration period of 45–60 min, the corneal thickness was measured again to determine the zero reference value for corneal swelling calculations.

At time $t = 0$ (i.e. immediately after the zero reference measurement) the test substance was applied to the eye. For this purpose, the clamp holding the eye was placed on a paper tissue outside the chamber with the cornea facing upwards. Because the surface area of the chicken cornea is only one-third that of the rabbit cornea, the standard dosing volume of 0.1 ml or 0.1 g of the test substance to be applied was adjusted by one-third to 0.03 ml or 0.03 g. Liquid materials were applied in amounts of 0.03 ml from a micropipette (Nichiryo Co. Ltd, model 8100, Tokyo, Japan), in such a way that the entire surface of the cornea was bathed with the test substance. Solids, ground to a fine powder if necessary, were applied in an amount of 30 mg by powdering the entire surface of the cornea. After a total exposure period of 10 sec, the corneal surface was rinsed thoroughly with 20 ml isotonic saline of ambient temperature and the eye in the holder was returned to its chamber. This procedure was repeated for each test eye. Each test compound was tested on five out of the six eyes; the sixth eye was treated in a similar way with isotonic saline only and served as a control. The control eye and test eyes were examined at 30, 75, 120, 180 and 240 min after treatment, using the criteria and scoring system described below. All examinations were carried out with the slit-lamp microscope.

Criteria and scoring system

Corneal swelling. Corneal swelling, expressed as a percentage, was calculated according to the following formula:

$$\frac{\text{corneal thickness at time } t - \text{corneal thickness at time } t = 0}{\text{corneal thickness at } t = 0} \times 100\%$$

The percentage of swelling was calculated for each test eye at the observation time points of 30, 75, 120, 180 and 240 min. On the basis of both the mean percentage of corneal swelling of all test eyes at each time point and the time of the occurrence of such swelling, an overall category score was given as indicated in Table 1.

Corneal opacity. For each eye, at each of the observation time points, corneal opacity was scored from 0 to 4, using the following criteria: 0 = no opacity; 1 = scattered or diffuse areas, details of iris clearly visible; 2 = easily discernible translucent area, details of iris slightly obscured; 3 = severe corneal opacity, no details of iris visible, size of pupil barely discernible; 4 = complete corneal opacity, iris invisible. Based on the highest mean score for corneal opacity, as observed at any of the time points, an overall category score was given as indicated in Table 2.

Fluorescein retention. The fluorescein retention value of each test eye was scored from 0 to 3 at the observation time point of 30 min only, using the following criteria: 0 = no fluorescein retention; 1 = small number of cells retaining fluorescein; 2 = individual cells and areas of the cornea retaining fluorescein; 3 = large areas of the cornea retaining fluorescein. On the basis of the mean value, a category score was assigned, as indicated in Table 3.

Morphological effects. These include 'pitting' of corneal epithelial cells, 'loosening' of epithelium, 'roughening' of the corneal surface and 'sticking' of the test substance to the cornea. These findings could vary in severity and could be observed as a single morphological effect or in combination. The

Table 1. Overall category scores for corneal swelling

Mean corneal swelling (%)	Time swelling first occurred	Category
−5–5	At any time point	I
6–12	At any time point	II
13–18	At $t = 120$, 180 and/or 240 min	II
	At $t = 30$ and/or 75 min	III
19–26	At any time point	III
27–32	At $t = 120$, 180 and/or 240 min	III
	At $t = 30$ and/or 75 min	IV
>32	At any time point	IV

Table 2. Overall category scores for corneal opacity

Highest mean opacity score observed during test	Category
0.0–0.5	I
0.6–1.5	II
1.6–2.5	III
2.6–4.0	IV

classification of these findings was subject to interpretation by the investigator.

Recording of morphological effects was primarily to support possible future histological examination of the eyes and was used only for overall classification of *ex vivo* eye irritancy when loosening of epithelium would be observed.

Assessment of ex vivo irritancy grades

For each test compound, the ultimate *ex vivo* irritancy classification was made by selecting one of the following grades: not, slightly, moderately, or severely irritating. These grades correspond to the following combinations of category scores as given for corneal swelling, opacity and fluorescein retention (Tables 1–3): not irritating = at most one category score of II; slightly irritating = at least two category scores of II or at most one category of III; moderately irritating = at least two category scores of III or at most one category of IV; and severely irritating = at least two category scores of IV.

Extrapolation from ex vivo results to EC classification

The classification of the irritancy *in vivo* of the 21 compounds was assessed according to Directive 83/467/EEC (EEC, 1983). Data on irritancy *in vivo* of the 21 selected chemicals were obtained from Botham *et al.* (1989). Additional data on two chemicals (1-butanol and 2-methoxyethanol) were derived from Prinsen (1991a,b). According to Directive 83/467/EEC, the grades of irritancy *in vivo* are divided into three classes, namely non-irritating (NI), irritating to eyes (R36) and risk of serious damage to eyes (R41).

RESULTS

Species comparison study

The supply of cows at the selected slaughterhouse was rather irregular, and the age and condition of the animals varied considerably, from old milk cattle to injured or sick animals, which had to be killed because of their moribund condition. Despite precise instructions to the personnel, rough handling of the eyes during dissection could not be avoided.

Table 3. Overall category scores for fluorescein retention

Mean fluorescein retention score at $t = 30$ min	Category
0.0–0.5	I
0.6–1.5	II
1.6–2.5	III
2.6–3.0	IV

The intensive working pressure was another unfavourable factor. On examination of the enucleated eyes at the testing facility, it appeared that none of the corneas met the quality standards of the laboratory, in that none was completely undamaged. From these results it was concluded that obtaining a sufficient number of bovine eyes of standardized quality and with intact corneas would be very laborious, under the prevailing conditions of the slaughterhouse; therefore the bovine eye was considered not suitable for routine testing in the EET.

The supply of pigs at the slaughterhouses was very constant and all animals were young adult pigs. At the slaughterhouse in Hilversum, the very restricted working space at the bleeding location and the high speed of the process line hampered careful dissection of the eyes. At the slaughterhouse in Bussum, the process line between the bleeding location and the hot water bath was one straight line with sufficient working space; in addition, the speed of the process line allowed correct dissection of the eyes by the investigator. The dimensions of the porcine eye are similar to those of the rabbit eye; no difficulties were encountered, therefore, in processing the eyes. On examination of the enucleated eyes, it appeared that eyes with an intact cornea and with acceptable fluorescein staining could indeed be obtained. However, corneal measurements confirmed that the cornea of the untreated porcine eye is relatively thick (approximately 0.90 instrument units) when compared with the enucleated rabbit eye (approximately 0.60 units). The value of 0.90 was already in the top range of the measurement device. It was anticipated, therefore, that examination of corneal swelling after treatment with severe eye irritants would be difficult. Treatment of the isolated eyes with 1% aqueous sodium hydroxide, a severe irritant, did indeed show that the top limit on the scale was already reached within 75 min of treatment, representing a swelling of about 30%. Previous studies with 1% NaOH (Burton *et al.*, 1981; Koëter and Prinsen, 1985) showed that a swelling of more than 30% can easily occur; because of these practical constraints, the pig was considered less suitable as an eye donor for the EET.

At the selected poultry slaughterhouse, approximately 45,000 chickens were processed daily and the supply of animals was constant. Furthermore, the birds were all of the same strain and of approximately the same age and weight. The high speed of the process line (approximately 100 chickens per minute) made it impossible to collect suitable eyes; a sufficient number of heads was therefore collected and transported to the testing facility. This could easily be done, because almost immediately after decapitation the eyelids of chickens close and thus prevent corneal damage during transportation. Dissection of the eyes from the

heads, without damaging the cornea, was relatively easy. Like the porcine eye, the dimensions of the chicken eye do not differ significantly from those of the rabbit. From 12 heads, at least 12 undamaged eyes could be obtained. Six eyes with intact cornea and with acceptable fluorescein staining could easily be selected. The corneal thickness of the untreated chicken eye was about 0.60 units. This allowed for corneal thickness measurements of compounds ranging from non-irritant to severely irritant. Preliminary investigations with 1% aqueous sodium hydroxide showed a good correlation with the results of the rabbit EET.

The preliminary observations with the three species of slaughter animals clearly indicated that the chicken offered the best technical possibilities. Therefore, this species was selected as the most suitable eye donor for the EET with slaughter eyes. Consequently, the comparative study with 21 reference compounds was carried out with the enucleated eyes of chickens.

Comparative study

Mean values for corneal swelling, corneal opacity and fluorescein retention obtained with each of the 21 compounds are shown in Table 4. The mean corneal swelling percentages ranged from −1% (minor shrinking) to 60%, which is within the range found in previous studies with enucleated rabbit eyes carried out in our Institute (Koëter and Prinsen, 1985; Prinsen and Koëter, 1990). Control eyes (one per compound) never showed any unusual swelling (between −4% and +4%). Slight shrinking of the cornea is a common finding when testing non-irritant materials or maintaining enucleated control eyes in the superfusion apparatus. Mean corneal opacity scores observed in this study ranged from 0.0 to 3.1; mean fluorescein retention scores ranged from 0.0 to 3.0.

Table 4. Mean values for corneal swelling, corneal opacity and fluorescein retention obtained from each test with five test eyes exposed to different test compounds* at various time intervals, *t* (min after treatment)

	Mean percentage corneal swelling at *t*					Mean corneal opacity at *t*					Mean fluorescein retention at *t*
Test compound	30	75	120	180	240	30	75	120	180	240	30
Acetic acid	12	20	25	25	31	2.0	2.5	2.6	2.2	2.2	3.0
	1†	2	2	2	4	0.0	0.0	0.1	0.2	0.2	0.0
Brij 35	1	2	2	4	5	0.0	0.0	0.0	0.0	0.0	0.9
	2	1	1	2	2	0.0	0.0	0.0	0.0	0.0	0.1
Benzalkonium chloride‡	15	17	23	30	40	2.8	3.0	3.0	3.0	3.0	3.0
	1	2	2	3	2	0.2	0.0	0.0	0.0	0.0	0.0
Dimethyl sulphoxide	1	1	3	4	3	0.5	0.5	0.5	0.5	0.5	1.0
	1	1	1	1	1	0.0	0.0	0.0	0.0	0.0	0.0
Sodium fluorescein	−1§	0	−1	−1	−1	0.0	0.0	0.0	0.0	0.0	0.1
	0	1	0	1	1	0.0	0.0	0.0	0.0	0.0	0.1
Glycerol	2	2	2	2	4	0.2	0.2	0.3	0.3	0.4	0.5
	1	2	1	1	2	0.1	0.1	0.1	0.1	0.1	0.0
Triacetin	4	2	4	2	1	0.1	0.1	0.1	0.1	0.1	0.5
	1	1	1	1	1	0.1	0.1	0.1	0.1	0.1	0.0
Mercury(II) chloride	15	41	49	52	55	2.6	2.6	2.6	3.1	3.1	2.0
	3	12	10	11	10	0.1	0.1	0.1	0.1	0.1	0.0
Silver(I) nitrate	4	9	11	11	12	0.8	0.8	1.0	1.0	1.0	1.0
	1	2	1	1	1	0.0	0.0	0.0	0.0	0.0	0.0
Sodium hydroxide‡	11	21	36	51	60	1.0	1.6	2.0	2.5	3.0	3.0
	0	1	2	1	1	0.0	0.1	0.0	0.0	0.0	0.0
Toluene	2	2	2	3	4	0.5	0.5	1.0	1.0	1.4	1.1
	0	2	2	2	2	0.0	0.0	0.0	0.0	0.1	0.2
Triethanolamine	2	1	2	3	4	0.5	0.6	0.6	0.6	0.7	0.9
	1	1	2	2	1	0.0	0.1	0.1	0.1	0.1	0.1
n-Hexane	1	0	0	0	−1	0.0	0.0	0.0	0.0	0.0	0.5
	1	1	1	2	2	0.0	0.0	0.0	0.0	0.0	0.0
Chloroform	7	8	12	18	21	1.0	1.0	1.0	1.0	1.0	2.5
	1	1	2	2	1	0.0	0.0	0.0	0.0	0.0	0.0
2-Methoxy ethanol	9	12	12	15	18	1.0	1.8	1.9	1.9	2.0	2.0
	1	1	1	1	2	0.0	0.2	0.1	0.1	0.0	0.2
1-Butanol‡	17	35	47	54	53	0.9	1.4	1.7	1.9	2.0	2.9
	1	3	5	5	4	0.1	0.2	0.1	0.1	0.0	0.1
Acetaldehyde	12	16	19	21	24	0.5	0.6	0.8	1.2	1.4	2.0
	1	1	2	2	2	0.0	0.1	0.2	0.3	0.2	0.0
2-Butoxyethyl acetate	3	4	5	4	4	0.5	0.5	1.0	1.0	1.0	1.0
	1	2	3	3	3	0.0	0.0	0.0	0.0	0.0	0.0
SDS‡	8	14	17	20	22	0.2	0.5	0.7	0.7	1.0	0.8
	1	2	2	2	3	0.1	0.2	0.1	0.1	0.0	0.1
Dibutyltin dichloride	14	18	22	28	34	2.5	2.5	2.5	2.5	2.5	3.0
	2	2	2	2	2	0.0	0.0	0.0	0.0	0.0	0.0
Tributyltin chloride	12	18	25	35	48	1.9	2.0	2.5	2.5	2.5	3.0
	1	1	2	2	1	0.1	0.0	0.0	0.0	0.0	0.0

*Chemicals were tested undiluted with the exception of acetic acid, silver(I)nitrate, sodium fluorescein and sodium hydroxide, which were tested as 10, 3, 20 and 1% dilutions in demineralized water, respectively.
†Standard error of the mean.
‡Loosening of epithelium was observed in one or more eyes.
§Negative value represents shrinking of the cornea.

Table 5. Categories for corneal swelling, corneal opacity, and fluorescein retention on the basis of the highest mean scores, subsequent CEET irritancy classification *ex vivo* on the basis of these overall categories and EC classification* of 21 compounds

Test compound	Corneal reaction†		Fluorescein retention	Classification	
	Swelling	Opacity		*ex vivo*	EC‡
Acetic acid	III	IV	IV	Severe	R41
Brij 35	I	I	II	non	NI
Benzalkonium chloride	IV	IV	IV	Severe	R41
Dimethyl sulphoxide	I	I	II	Non	NI
Sodium fluorescein	I	I	I	Non	NI
Glycerol	I	I	I	Non	NI
Triacetin	I	I	I	Non	NI
Mercury(II)chloride	IV	IV	III	Severe	R41
Silver(I)nitrate	II	II	II	Slight	NI§
Sodium hydroxide	IV	IV	IV	Severe	R41
Toluene	I	II	II	Slight	NI
Triethanolamine	I	II	II	Slight	NI
n-Hexane	I	I	I	Non	NI
Chloroform	III	II	III	Moderate	R36
2-Methoxy ethanol	II	III	III	Moderate	R36‖
1-Butanol	IV	III	IV	Severe	R41
Acetaldehyde	III	II	III	Moderate	R36‖
2-Butoxyethyl acetate	I	II	II	Slight	NI
SDS	III	II	II	Severe¶	R41
Dibutylin dichloride	IV	III	IV	Severe	R41
Tributyltin chloride	IV	III	IV	Severe	R41

*Data from Botham *et al.* (1989), except for 2-methoxy ethanol and 1-butanol (data from Prinsen, 1991a,b).
†I = no effect; II = slight effect; III = moderate effect; IV = severe effect.
‡NI = non-irritating; R36 = irritating to eyes; R41 = risk of serious damage to eyes.
§Borderline case between NI and R36.
‖Borderline case between R36 and R41.
¶Upgraded from R36 to R41 because of a loosening of epithelium.

Other effects observed consisted of loosening of epithelium (caused by benzalkonium chloride, NaOH, *n*-butanol and SDS), pitting (caused by acetic acid, toluene, SDS, triethanolamine, 1-butanol and 2-methoxyethanol), adherence of the test substance to the cornea or precipitation of the substance on the cornea [caused by dibutyltin dichloride and by silver nitrate (the latter probably formed silver chloride in combination with the saline dropped onto the eye)] and small vesicles in or on the cornea (caused by tributyltin chloride). In the case of 2-butoxy ethyl acetate an unattached layer was observed on the cornea of all test eyes approximately 3 hr after treatment. A possible explanation for this phenomenon could not be found, but determination of fluorescein retention showed no additional damage when compared with the 30-min determination. Loosening of epithelium was observed with four compounds, of which three (benzalkonium chloride, NaOH and 1-butanol) would already be classified as severe irritants on the basis of the results observed; the fourth compound (SDS) would be classified as (moderately) irritating on the basis of corneal swelling, opacity and fluorescein retention. However, because loosening of epithelium is a severe effect, it was considered justified to upgrade the classification of this compound to severely irritating. The significance of the other effects was not fully understood and the effects were therefore not included in the assessment of the irritancy classification.

Assessment of ex vivo irritancy categories

On the basis of the highest mean values for corneal swelling, corneal opacity and fluorescein retention, the respective categories for these parameters were assigned to each of the 21 compounds (Table 5), using the criteria mentioned in Tables 1–3. The EC classifications of the 21 compounds are also presented in Table 5. When compared with the EC classification, the CEET correctly classified as moderate or severe irritants each of the compounds that, in the EC, must be labelled R36 or R41 on the basis of *in vivo* studies. Four compounds [silver(I) nitrate, toluene, triethanolamine and *n*-hexane], which on the basis of the EC classification are considered as non-irritant, appeared to be slightly irritating in the CEET.

DISCUSSION

In the past, the enucleated eye test has been recognized as a valuable alternative to the Draize eye irritation test with rabbits, because it represents a test system nearest to the *in vivo* test method. As with the Draize rabbit test, the EET can be applied to substances with various physical characteristics, including liquids and solids, without dilution. In addition, no complicated techniques are necessary to perform the test, and the use of 'whole' eyes with the cornea *in situ* enables examination of the reaction of 'intact' corneal tissue to chemical insult and allows for histological examination of this tissue. In the EET, three parameters are measured as indicative of possible adverse eye effects, namely corneal thickness (expressed as corneal swelling), corneal opacity and fluorescein retention. The measurement of corneal swelling in this assay provides an objective parameter that enables the investigator to determine the

damaging effects of test materials very precisely, in contrast to the conventional rabbit test in which only subjective, gross observations are made. In combination with the detailed observation of corneal opacity and fluorescein retention by means of the slit-lamp microscope, a reliable evaluation of the eye irritation potential of test materials can be achieved.

So far, the rabbit has been used as the eye donor for the EET; in our laboratory only those rabbits were used that had already participated in other studies, such as the acute dermal irritation study, and would therefore have been killed anyway. However, the stock of such animals is irregular and not easily accessible for many laboratories. To avoid completely the killing of any laboratory animals for the EET, the use of slaughter animals as eye donors, such as cattle, pigs or chickens, would be a good alternative. From these donors, the chicken appeared to be the best candidate, particularly in view of the numerous chickens processed daily and because of the ease of obtaining eyes with corneas suitable for testing.

As all irritating compounds examined in this comparative study with chicken eyes were correctly classified when compared with the EC classification based on *in vivo* studies, it seems justified to conclude that this *ex vivo* test system is reliable and accurate in the assessment of eye irritation potential of test materials without the use of laboratory animals. In addition, although the CEET does not predict damage to conjunctival tissue, this deficiency did not reduce the sensitivity with which it predicted ocular irritancy. The relationship between conjunctival reactions and corneal swelling, simultaneously assessed during studies *in vivo*, has already been reported by Burton (1971).

Four compounds were classified as slightly irritant on the basis of the CEET study, whereas, according to the EC classification, they are considered as non-irritants. However, whereas the EC classification distinguishes three categories, the CEET recognizes four categories and discriminates also between slight and moderate irritants. As the scale for eye irritancy should be considered a sliding scale rather than one of three clear-cut steps, the cut-off point for non-irritancy in the CEET is obviously somewhat lower than that of the EC classification, which is based on the Draize test. One could argue about the validity of either classification, since data from human exposure are necessary for any justification and such data are not sufficiently available. Therefore, and because one of the four compounds in question [silver(I) nitrate] was already classified as a borderline between non-irritant and irritating (R36), the CEET can be considered as a sensitive, but not over-sensitive, means of predicting the eye irritancy potential of all types of compound.

In their latest updates of guidelines on eye irritancy testing, the EEC and the OECD have already recommended the use of alternative *ex vivo*/*in vitro* test systems for prescreening or positive identification of strong eye irritants; the CEET can make a valuable contribution to such an approach. In order that this test may also be used for the identification of non-irritants, where the CEET gives a negative result it is recommended that an additional single-animal test be conducted [OECD guideline no. 405 (OECD, 1987)] instead of a full three-animal study, to confirm non-irritancy of test materials. By this approach, animal suffering and animal use will be reduced to a minimum and, in time, a vast amount of data *in vivo* and *ex vivo* on non-irritant or slightly irritating materials will become available, allowing further evaluation and interpretation of these very mild effects.

Acknowledgements—We thank the Platform Alternatieven voor Dierproeven (Dutch society providing funds for the research of alternative methods for animal testing).

REFERENCES

Botham J. W., McKillop C. M. and Purchase I. F. H. (1989) Acute and topical toxicity profiles for substances involved in the *in vitro* validation of the *in vivo* ocular irritation model. In *Collaborative Study on the Evaluation of Alternative Methods to the Eye Irritation Test. EC Document XI/632/91, V/E/1/131/91.* Commission of the European Communities.

Burton A. B. G. (1971) A method for the objective assessment of eye irritation. *Food and Cosmetics Toxicology* **10**, 209–217.

Burton A. B. G., York M. and Lawrence R. S. (1981) The *in vitro* assessment of severe irritants. *Food and Cosmetics Toxicology* **19**, 471–480.

CEC (1991) Collaborative study on the evaluation of alternative methods to the eye irritation test. EC Document XI/632/91, V/E/1/131/91. Commission of the European Communities.

EEC (1983), EEC-Directive 83/467/EEC. *Official Journal of the European Communities*, L257. Vol. 26. pp. 19–20.

Koëter H. B. W. M. and Prinsen M. K. (1985) Introduction of an *in vitro* eye irritation test as a possible contribution to the reduction of the number of animals in toxicity testing. TNO Report V85.188/140322, May 1985, Dept. of Biological Toxicology, TNO Toxicology and Nutrition Institute, Zeist, The Netherlands. (Available on request.)

Koëter H. B. W. M. and Prinsen M. K. (1985) Comparison of *in vivo* and *in vitro* eye irritancy test systems: a study with 34 substances. In *Alternative Methods in Toxicology*. Vol. 3, Chap. A9. Edited by A. M. Goldberg. Mary Ann Liebert, New York.

OECD (1987) Guideline no. 405. Acute eye irritation/corrosion, second updated version. Organisation for Economic Co-operation and Development, Paris.

OECD (1991) Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1: The OECD Principles of Good Laboratory Practice. Organisation for Economic Co-operation and Development, Paris.

Price J. B. and Andrews I. J. (1985) The *in vitro* assessment of eye irritancy using isolated eyes. *Food and Chemical Toxicology* **23**, 313–315.

Prinsen M. K. and Koëter H. B. W. M. (1990) EC-Validation study on alternatives to the Draize eye irritation test. Pilot Interlaboratory comparison of the enucleated eye test. TNO Report V89.464, April 1990, Dept. of Biological Toxicology, TNO Toxicology and Nutrition Institute, Zeist, The Netherlands. (Available on request.)

Prinsen M. K. (1991a) Acute eye irritation/corrosion study with 1-butanol in albino rabbits. TNO-report V91.187/210069, November 1991. Dept. of Biological Toxicology, TNO Toxicology and Nutrition Institute, Zeist, The Netherlands. (Available on request.)

Prinsen M. K. (1991b) Acute eye irritation/corrosion study with 2-methoxyethanol in albino rabbits. TNO-report V91.188/210069, November 1991. Dept. of Biological Toxicology, TNO Toxicology and Nutrition Institute, Zeist, The Netherlands. (Available on request.)

The Chicken Enucleated Eye Test as a prescreen in routine toxicity testing

M.K. Prinsen.
The Chicken Enucleated Eye Test (CEET): A practical (pre)screen for the assessment of eye irritation/corrosion potential of test materials.

Food and Chemical Toxicology Vol. 34, No. 3, pp. 291-296, 1996

4

Pergamon

0278-6915(95)00115-8

Fd Chem. Toxic. Vol. 34, No. 3, pp. 291–296, 1996
Elsevier Science Ltd
Printed in Great Britain
0278-6915/96 $15.00 + 0.00

The Chicken Enucleated Eye Test (CEET): a Practical (Pre)Screen for the Assessment of Eye Irritation/Corrosion Potential of Test Materials

M. K. PRINSEN

TNO Nutrition and Food Institute, Toxicology Division, PO Box 360, 3700 AJ Zeist, The Netherlands

(Accepted 13 October 1995)

Abstract—The enucleated eye test with chicken eyes (CEET) obtained from an abattoir proved to be a valuable and practical alternative for the 'traditional' enucleated eye test with eyes of laboratory rabbits. Since 1992, the CEET has been incorporated in standard contract toxicity testing at the Toxicology Division of the TNO Nutrition and Food Research Institute as a (pre)screen for the Draize eye test with rabbits. The results of the first 44 compounds tested showed excellent correlation with the *in vivo* results. The CEET identified non-irritating or severely irritating compounds, and predicted (slightly to moderately) irritating compounds. Statistical analysis of the CEET and the rabbit *in vivo* scores showed high linear correlations between the critical values of both tests and confirmed the relevance of this assay with respect to ocular effects. In general, tiered *in vitro*/*in vivo* testing was considered a meaningful approach for further validation of alternative methods and for reducing the use of suffering of laboratory animals to a minimum. Tiered testing of compounds in cases of eye irritation hazard assessment should be incorporated in the legislation of the European Community.

INTRODUCTION

The enucleated eye test with isolated eyes of rabbits was first introduced by Burton *et al.* in 1981. Parameters such as corneal thickness, corneal opacity and fluorescein retention were used for disclosing possible adverse eye effects of test materials. Corneal thickness is recognized as a reliable and objective parameter for the assessment of corneal injury (Burton, 1971). Slaughterhouse waste tissue was investigated as a source of eyes in order that the use of laboratory rabbits as eye donors could be abandoned. Of possible eye donor species such as the cow, pig and chicken, the latter was considered the most suitable. In a comparative study with 21 chemicals the chicken enucleated eye test (CEET) was shown to be reliable and accurate in assessing the eye irritation potential of test materials (Prinsen and Koëter, 1994). In their latest updates of the guidelines on eye irritancy testing both the European Community (EC) and the Organisation for Economic Cooperation and Development (OECD) allow the use of alternative *in vitro* test systems for screening of severe eye irritants, to reduce animal use and suffering. Therefore, a programme was initiated in the TNO Nutrition and Food Institute to introduce the CEET in contract toxicity testing as a (pre)screen for the standard Draize eye irritation test. Depending on the results of the CEET, which are available within 6 hr after the start of the test, the sponsor could decide whether to confirm non-irritancy, irritancy or even severe irritancy in the *in vivo* rabbit eye irritation test. Preliminary investigation on the performance of this kind of tiered testing showed very promising results (Prinsen, 1995). Therefore, it was decided to permanently use the CEET for (pre)screening at TNO. This paper reports the findings of the first 44 compounds tested in a tiered fashion. All tests (CEET and the *in vivo* rabbit eye test) were carried out according to the OECD principles of good laboratory practice (OECD, 1991) to obtain maximum acceptance. In the first instance, reporting of the results was done either in separate reports or, preferably, combined into one report to allow regulatory authorities to become familiar with the test method. From 1994 only combined reporting was performed.

Abbreviations: CEET = chicken enucleated eye test; EC = European Community; OECD = Organisation for Economic Cooperation and Development.

MATERIALS AND METHODS

Materials

44 test materials received from 1992 to 1994, were allowed for testing in both test systems by various international sponsors. The compounds consisted of

Table 1. *In vivo* scores used for X/Y comparison with CEET scores*

1-hr score	= mean score after 1 hr (1 hr)
24-hr score	= mean score after 24 hr (24 hr)
1–72 hr score	= mean score over the 1–72 hr period (1–72 hr)
Total eye score	= all eye scores obtained added (Total score)
MAS score	= maximum average score (MAS)
Opacity score	= total scores for opacity (OP)
Area score	= total scores for the area of opacity (AS)
Iris score	= total scores for iritis (IS)
Redness score	= total scores for redness of the conjunctivae (RS)
Swelling score	= total score for swelling of the conjunctivae (SS)
Discharge score	= total scores for discharge of the conjunctivae (DS)
Conjunctival score	= total scores for all effects of the conjunctivae (TC)
Iris and opacity score	= total scores for iritis and opacity (I/O)
Recovery score	= number of days to recovery (DR)

**In vivo* scores recommended by the Interagency Regulatory Alternatives Group of the United States (workshop on eye irritation testing: practical applications of non-whole animal alternatives, Washington, DC, November 1993) and calculated with the Draize grading system (1944).

33 liquids, nine solids, one paste and one gel. As information on part of the test materials was classified by the sponsor as confidential, most of the 44 compounds are only generally described. The classes of compounds tested represent industrial chemicals, pesticides, detergents, commercial formulations and foodstuffs. Exactly the same batch of the compound was used for both tests.

Experimental design

Chicken enucleated eye test. The CEET, its scoring and (EC) classification have been fully described in an earlier publication on the CEET (Prinsen and Koëter, 1994). Heads of spring chickens were collected at a local abattoir. Enucleated eyes, within 2 hr of death, were placed in a superfusion apparatus, which kept the eyes in good condition for at least 6 hr

Table 2. *In vivo* scores of the 39 compounds used for data analyses*

	In vivo scores													
Test compound	1 hr	24 hr	1–72 hr	Total score	MAS	OS	AS	IS	I/O	RS	SS	DS	TC	DR
1. Formulation	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2. Formulation	37	31	36	463	44	18	32	10	60	31	23	32	86	21
3. Pesticide	13	5	7	78	13	0	0	2	2	15	12	6	33	7
4. Detergent	18	12	13	157	18	0	0	3	3	22	20	22	64	21
5. Silicone powder	5	1	2	18	5	0	0	0	0	2	3	1	6	2
6. Lubricant	2	4	2	24	4	0	0	0	0	9	3	0	12	3
7. Ink	3	0	1	8	3	0	0	0	0	3	1	0	4	1
8. Ink	2	0	1	6	2	0	0	0	0	3	0	0	3	1
9. Paint	13	6	8	79	13	0	0	2	2	14	15	11	40	7
10. Silicone powder	6	2	2	24	6	0	0	0	0	6	3	3	12	2
11. Sodium *p*-styrene sulfonate	23	33	21	289	33	10	17	6	33	27	14	17	58	17
12. Formulation	36	37	49	1055	68	19	46	7	72	29	35	36	100	>21
13. Pesticide	5	1	2	18	5	0	0	0	0	5	2	2	9	3
14. Polydisaccharide (14.5%)	2	0	0	5	0	0	0	0	0	1	0	0	1	1
15. Polydisaccharide (50%)	8	0	2	248	8	0	0	0	0	3	3	6	12	1
16. Liquid nylon product	6	1	1	30	4	0	0	0	0	11	3	1	15	3
17. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
18. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
19. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
20. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
21. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
22. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
23. Solvent	1	0	0	2	1	0	0	0	0	1	0	0	1	1
24. Solvent	1	0	0	2	1	0	0	0	0	1	0	0	1	1
25. Solvent	2	1	1	8	2	0	0	0	0	4	0	0	4	2
26. Ink	1	0	0	4	1	0	0	0	0	2	0	0	2	1
27. Thermal paper coating	11	23	17	206	24	6	21	4	31	22	16	8	46	7
28. Toilet cleaner	15	7	8	121	15	0	0	3	3	17	18	8	43	28
29. Toilet cleaner	29	30	20	241	32	5	20	4	29	25	15	19	59	14
30. Pesticide	8	1	2	57	8	0	0	3	3	12	4	5	21	3
31. Sulfur	6	3	3	60	6	0	0	0	0	17	7	6	30	3
32. Ink	0	4	2	52	4	0	0	0	0	14	12	0	26	3
33. Thermal paper coating	4	0	1	11	4	0	0	1	1	3	0	0	3	1
34. Detergent	40	48	52	819	68	21	43	12	76	28	35	36	99	30
35. Propyl-lactate	61	63	63	1090	63	24	48	12	84	33	36	36	105	30
36. Ethylhexyl-lactate	37	50	37	534	52	18	35	9	62	25	17	23	65	17
37. Pesticide	10	11	25	101	12	0	0	1	1	19	16	13	48	7
38. Solvent	2	0	1	6	2	0	0	0	0	3	0	0	3	1
39. Detergent	1	1	1	6	1	0	0	0	0	3	0	0	3	2

*Scores calculated with the Draize grading system (1944).

Table 3. CEET scores of the 39 compounds used for data analyses and the five compounds not tested *in vivo*

Test compound	CEET index	% swelling	Opacity (×20)	Fluorescein retention (×20)
1. Formulation	0	0	0	0
2. Formulation	118	24	40	54
3. Pesticide	21	3	6	12
4. Detergent	69	9	30	30
5. Silicone powder	0	0	0	0
6. Lubricant	1	1	0	0
7. Ink	18	2	0	16
8. Ink	5	3	0	2
9. Paint	41	5	10	26
10. Silicone powder	1	1	0	0
11. Sodium *p*-styrene sulfonate	85	19	26	40
12. Formulation	125	35	40	50
13. Pesticide	15	1	0	14
14. Polydisaccharide (14.5%)	8	2	0	6
15. Polydisaccharide (50%)	2	2	0	0
16. Liquid nylon product	1	1	0	0
17. Solvent	6	0	0	6
18. Solvent	0	0	0	0
19. Solvent	0	0	0	0
20. Solvent	19	3	6	10
21. Solvent	12	0	6	6
22. Solvent	10	0	6	4
23. Solvent	6	2	0	4
24. Solvent	7	3	0	4
25. Solvent	1	1	0	0
26. Ink	2	0	0	2
27. Thermal paper coating	41	9	12	20
28. Toilet cleaner	56	12	16	28
29. Toilet cleaner	57	11	20	26
30. Pesticide	57	7	20	30
31. Sulfur	5	1	0	4
32. Ink	37	7	10	20
33. Thermal paper coating	55	5	10	40
34. Detergent	105	25	40	40
35. Propyl-lactate	165	45	60	60
36. Ethylhexyl-lactate	98	18	40	40
37. Pesticide	65	15	20	30
38. Solvent	3	3	0	0
39. Detergent	24	4	10	10
40. Glycolbromoacetate form.	131	41	38	52
41. Amidosulfonic acid	180	46	80	54
42. Glycolbromoacetate 85%	156	36	60	60
43. Monobromoacetic acid	200	80	80	60
44. Didec.meth.amm.chl.*	169	39	70	60

*23% didecyldimethylammoniumchloride in propyl glycol.

as shown by the control eyes, that is, showing a corneal swelling of less than 5% and absence of corneal opacity and fluorescein retention. The test compound was applied to the cornea of five or three eyes in one single dose of either 30 μl (liquids, gels and pastes) or 30 mg (solids) for 10 sec. Before dosing each eye provided its own baseline values for the assessment of the corneal effects. Control eyes, one per test run, were used only for checking the experimental conditions. The reactions of the corneas, that is the occurrence of swelling, opacity or fluorescein retention of damaged epithelial cells, were examined up to 4 hr after treatment with a slit-lamp microscope.

In vivo rabbit eye test. All studies were conducted in accordance with the relevant OECD (1987) and EC (1992) guidelines on eye irritation testing. An amount of 0.1 ml of the test substance (solids as well as liquids) was instilled in the conjunctival sac of the right eye of each rabbit. After administration the upper and lower eye lid were carefully closed and subsequently held together for at least 1 sec before releasing, to prevent loss of material. The left eye remaining untreated served as a control. The reactions of the test eyes were judged at about 1, 24, 48 and 72 hr after treatment using the Draize scoring system (Draize, 1944). Residual eye effects were recorded at regular intervals, if necessary up to about 3 wk after treatment.

Evaluation of the results

On the basis of the maximum mean scores observed for corneal swelling, corneal opacity and fluorescein retention of damaged epithelial cells, the irritation potential of the compound was assessed using the CEET EC classification system (Prinsen and Koëter, 1994). Both classifications of the 44 compounds were compared after applying the EC criteria (EC, 1993) for the data.

Statistical analyses of the data were performed according to guidelines for the evaluation of eye irritation alternative tests; criteria for data submissions as recommended by the Interagency Regulatory Alternatives Group of the United States (workshop on eye irritation testing: practical applications of non-whole animal alternatives, Washington, DC, November 1993). These guidelines state the necessary steps in the evaluation of data to assess the utility of *in vitro* tests, such as defining the proposed context for use of the *in vitro* method, collection and collation of data, preparation of plots showing the relationship between the *in vitro* and *in vivo* data, statistical analysis and reporting. Following these directions, 14 different *in vivo* scores (Table 1) were calculated after applying the multipliers mentioned in the Draize grading system, namely a factor 5 for cornea and iris lesions and a factor 2 for conjunctivae lesions.

An irritation index was calculated with the results of the CEET which, in addition to the individual components (i.e. corneal swelling, corneal opacity and fluorescein retention), could be used for statistical comparison with the *in vivo* scores. The principle of this index is the adding of the maximum mean scores of the three parameters measured, namely corneal swelling, corneal opacity and fluorescein retention. The opacity and fluorescein scores are almost equally weighed in the index compared with that of the percent maximum swelling obtained in this assay, that is, about 60–80% at the TNO Institute. Therefore, the maximum mean score for corneal opacity (score 4) and fluorescein (score 3) obtained with a given compound were multiplied by a factor of 20. Thus, the irritation index could range from 0 to 200. X/Y plots with Pearson's correlation coefficient were made with these *in vitro* and *in vivo* data points.

Table 4. Comparison *of in vitro* and *in vivo* EC classification of the 44 compounds

	EC classification	
Test compound*	CEET†	*In vivo* rabbit‡
1. Formulation (li)	NI	NI
2. Formulation (li)	R36	R36
3. Pesticide (li)	NI	NI
4. Detergent (li)	NI/R36	NI§
5. Silicone powder (s)	NI	NI
6. Lubricant (gel)	NI	NI
7. Ink (li)	NI	NI
8. Ink (li)	NI	NI
9. Paint (li)	NI	NI
10. Silicone powder (s)	NI	NI
11. Sodium-*p*-styrene sulfonate (s)	R36	R36
12. Formulation (paste)	R36/R41	R41
13. Pesticide powder (s)	NI	NI
14. Polydisaccharide 14.5% (li)	NI	NI
15. Polydisaccharide 50% (li)	NI	NI
16. Liquid nylon product	NI	NI
17. Solvent (li)	NI	NI
18. Solvent (li)	NI	NI
19. Solvent (li)	NI	NI
20. Solvent (li)	NI	NI
21. Solvent (li)	NI	NI
22. Solvent (li)	NI	NI
23. Solvent (li)	NI	NI
24. Solvent (li)	NI	NI
25. Solvent (li)	NI	NI
26. Ink (li)	NI	NI
27. Thermal paper coating (li)	NI/R36	NI§
28. Toilet cleaner (li)	NI/R36	NI§
29. Toilet cleaner (li)	NI/R36	NI§
30. Pesticide (s)	NI/R36	NI
31. Sulfur (s)	NI	NI
32. Ink (li)	NI	NI
33. Thermal paper coating (li)	NI	NI
34. Detergent (li)	R36	R41
35. Propyl-lactate (li)	R41	R41
36. Ethylhexyl-lactate (li)	R36	R36
37. Pesticide (s)	NI/R36	NI
38. Solvent (li)	NI	NI
39. Detergent (li)	NI	NI
40. Glycolbromoacetate form. (li)	R41	Skin corrosive/R41
41. Amido sulfonic acid (s)	R41	Skin corrosive/R41
42. Glycolbromoacetate 85% (li)	R41	Skin corrosive/R41
43. Monobromoacetic acid (s)	R41	Skin corrosive/R41
44. Didec.meth.amm.chl. (li)	R41	Skin corrosive/R41

*li = liquid; s = solid.
†CEET EC classification scheme (Prinsen and Koëter, 1994).
‡Classification according to EC standards (EC, 1993).
§Borderline eye effects.
NI = not irritant; R36 = irritating to eyes; R41 = risk of serious damage to eyes.

Table 5. *R* values and 95% confidence intervals obtained after data analyses (X/Y plots with Pearson's correlation coefficient) with 39 compounds

	CEET scores			
In vivo scores	Index	% swelling	Opacity	Fluorescein
1 hr	0.92 (0.85–0.96)	0.93 (0.87–0.96)	0.94 (0.88–0.97)	0.84 (0.72–0.91)
24 hr	0.90 (0.81–0.94)	0.91 (0.84–0.95)	0.92 (0.84–0.96)	0.81 (0.66–0.90)
1–72 hr	0.92 (0.86–0.96)	0.96 (0.92–0.98)	0.93 (0.87–0.96)	0.83 (0.70–0.91)
Total score	0.87 (0.77–0.93)	0.94 (0.89–0.97)	0.88 (0.78–0.94)	0.77 (0.60–0.87)
MAS	0.91 (0.83–0.95)	0.92 (0.86–0.96)	0.91 (0.84–0.95)	0.83 (0.69–0.91)
OP	0.87 (0.77–0.93)	0.91 (0.84–0.95)	0.89 (0.79–0.94)	0.78 (0.62–0.88)
AS	0.86 (0.75–0.92)	0.91 (0.83–0.95)	0.87 (0.76–0.93)	0.77 (0.60–0.87)
IS	0.92 (0.86–0.96)	0.91 (0.83–0.95)	0.94 (0.88–0.97)	0.86 (0.74–0.92)
RS	0.88 (0.78–0.94)	0.85 (0.73–0.92)	0.88 (0.78–0.94)	0.84 (0.71–0.91)
SS	0.90 (0.82–0.95)	0.91 (0.83–0.95)	0.90 (0.81–0.95)	0.84 (0.71–0.91)
DS	0.92 (0.85–0.96)	0.92 (0.85–0.96)	0.94 (0.88–0.97)	0.85 (0.73–0.92)
DR	0.88 (0.79–0.94)	0.88 (0.79–0.94)	0.89 (0.79–0.94)	0.82 (0.68–0.90)
I/O	0.88 (0.78–0.94)	0.92 (0.84–0.96)	0.89 (0.80–0.94)	0.79 (0.64–0.89)
TC	0.92 (0.86–0.96)	0.92 (0.85–0.96)	0.93 (0.87–0.96)	0.86 (0.75–0.93)
Grand mean ± SEM	0.90 ± 0.01	0.91 ± 0.01	0.86 ± 0.02	0.82 ± 0.01

RESULTS

All 44 compounds regardless of their physical form could be assayed in the CEET without any difficulties. 39 of the 44 compounds were examined in both the *in vivo* rabbit eye test (Table 2) and the CEET (Table 3). Six compounds showed very severe effects in the CEET and were considered severely irritating and possibly corrosive to eyes. In five cases, it was decided not to proceed with the *in vivo* eye test. These five compounds were, however, examined for acute skin irritation/corrosion in rabbits, and all were found to be corrosive to skin. The results after applying the CEET EC classification system and *in vivo* rabbit EC classification are presented in Table 4. Only two compounds (nos 12 and 34) were slightly underestimated by the CEET, that is, moderately irritating instead of severely irritating, while one compound (no. 37) was slightly overestimated, being borderline irritating in the CEET against non-irritating in the rabbit test.

The correlation coefficients and their 95% confidence interval obtained after comparison of the *in vivo* scores with the CEET values are summarized in Table 5. The overall linear correlation of the CEET irritation index with the various *in vivo* scores was very satisfactory. Very high correlation coefficients (up to 0.92) were obtained with f.i. the 1-hr score, the 1–72 hr score and the maximum average score. Even higher correlation coefficients were obtained after comparison of the *in vivo* scores with the CEET swelling (%). Almost maximum correlation was found with the 1–72 hr score (0.96) and the total score (0.94). The correlation between CEET opacity scores and *in vivo* scores was overall slightly lower but still satisfactory. The comparison with the fluorescein retention scores showed a lower but also acceptable correlation.

DISCUSSION AND CONCLUSIONS

The enucleated eye test with rabbit eyes is a well known method used internationally at various laboratories as an in-house method for (pre)screening compounds for eye irritation (Kočter and Prinsen, 1985a,b; Price and Andrews, 1985; York *et al.*, 1994). The CEET can be considered a logical step in the further development of this alternative method, by abandoning the use of laboratory animals by using slaughterhouse waste material. The enucleated eye test method is expected to have a high potential for replacement of the *in vivo* Draize eye test, because it uses "an eye for an eye" and uses discriminative parameters that represent actual *in vivo* eye damage. Furthermore, the test method is very straightforward and all kinds of materials can be readily assayed. The results obtained with 44 compounds so far fully confirm these expectations. High correlations were obtained after comparing the *in vivo* tissue scores of the 39 compounds with the CEET irritation index scores. The highest correlation was found after comparison of the percent corneal swelling and several individual *in vivo* scores. Lower but still satisfactory correlation was found between the *in vivo* scores and the corneal opacity or fluorescein retention scores of the CEET. The lower correlation with the latter two CEET parameters can be explained by the more crude and less objective assessment and scoring of these parameters compared with the assessment of corneal swelling. As about two-thirds of the 39 compounds were non-irritants, the distribution of the data is somewhat skewed to the lower part of the scale. The confidence limits, however, still show acceptable intervals for the majority of correlations. On the basis of the correlation found between the tissue scores of both test systems, the author feels that the use of the CEET as a robust and practical (pre)screen for the Draize rabbit eye test is confirmed and justified. The results after applying the CEET (EC) classification show its usefulness and accuracy in hazard classification according to the EC criteria.

In practice, the majority of compounds offered for hazard identification in contract research are non-irritants. Because the non-irritant and severely irritant compounds were correctly identified, the CEET

could be very useful for the identification of both severe and non-irritants without proceeding to the actual animal test. The mild to moderate irritants, generally showing the highest sensitivity to inter- and intra-laboratory variability, still need to be confirmed in the rabbit eye test. For hazard identification, in general, the author recommends that only compounds not expected to cause severe eye reactions as evaluated by an alternative non-animal test be assessed by the *in vivo* Draize eye test. Combined reporting of these results will make regulatory authorities more familiar with the performance of the alternative methods used. Because regulatory authorities still consider the rabbit test as 'the golden standard', for the time being, the *in vivo* results would still overrule the *in vitro* results in cases of clear discrepancies. In The Netherlands, the CEET has already been advised by regulatory authorities to check the corrosive properties of compounds suspected of corrosivity by sponsors on the basis of their own in-house experiences.

This kind of tiered eye irritation testing has been performed voluntarily at the TNO Institute since 1992 and is highly appreciated by sponsors, animal welfare organizations and regulatory bodies. Once incorporated in the routine toxicology screening of test materials, the costs appear to be low and insignificant against the benefit to the animals. The author believes that using the CEET reduces the suffering of rabbits in eye irritation testing to a minimum. This approach, however, is not intended to obstruct further optimization/modification of alternative methods and more mechanistic research or even the development of new alternatives. In contrast, it should stimulate these options.

REFERENCES

Burton A. B. G. (1971) A method for the objective assessment of eye irritation. *Food and Cosmetics Toxicology* **10**, 209–217.

Burton A. B. G., York M. and Lawrence R. S. (1981) The *in vitro* assessment of severe irritants. *Food and Cosmetics Toxicology* **19**, 471–480.

Draize J. H., Woodard G. and Calvery H. O. (1944) Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *Journal of Pharmacology and Experimental Therapeutics* **82**, 377–390.

EC (1992) Commission Directive 92/69/EEC, *Official Journal of the European Communities*, L383A, B.5. acute toxicity (eye irritation), 29 December.

EC Standards (1993) EC Directive 93/21/EC, *Official Journal of the European Communities*, L 110A, 4 May.

Koëter H. B. W. M. and Prinsen M. K. (1985a) Introduction of an *in vitro* eye irritation test as a possible contribution to the reduction of the number of animals in toxicity testing. TNO report V85.188/140322, May, Dept of Biological Toxicology, TNO Toxicology and Nutrition Institute, Zeist, The Netherlands (available on request).

Koëter, H. B. W. M. and Prinsen M. K. (1985b) Comparison of *in vivo* and *in vitro* eye irritancy test systems: a study with 34 substances, *Alternative Methods in Toxicology*. Vol. 3, chapter A9. Mary Ann Liebert, Inc., New York.

OECD (1987) Guideline no. 405 acute eye irritation/corrosion, 2nd updated version.

OECD (1991) *Series on Principles of Good Laboratory Practice and Compliance Monitoring, Number 1: The OECD Principles of Good Laboratory Practice*. OECD, Paris.

Price J. B. and Andrews I. J. (1985) The *in vitro* assessment of eye irritancy using isolated eyes. *Food and Chemical Toxicology* **23**, 313–315.

Prinsen M. K. (1995) The chicken enucleated eye test (CEET): a screen for the assessment of eye irritation/corrosion potential of compounds (results of 20 compounds). *Proceedings of the World Congress on Alternatives and Animal Use*. Edited by A. M. Goldberg. Life Sciences Book, New York.

Prinsen M. K. and Koëter H. B. W. M. (1994) Justification of the enucleated eye test with eyes of slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits. *Food and Chemical Toxicology* **31**, 69–76.

York M., Wilson A. P. and Newsome C. S. (1994) The classification of soluble silicates for eye hazard using the enucleated rabbit eye test. *In Vitro Toxicology* **8**, 1265–1269.

The Isolated Chicken Eye Test as a stand-alone test

K. Schutte, M.K. Prinsen, P.M. McNamee, R. Roggeband.
The isolated chicken eye test as a suitable *in vitro* method for determining the eye irritation potential of household cleaning products.

Regulatory Toxicology and Pharmacology 54 (2009) 272-281

5

Contents lists available at ScienceDirect

Regulatory Toxicology and Pharmacology

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

The isolated chicken eye test as a suitable *in vitro* method for determining the eye irritation potential of household cleaning products

K. Schutte [a,*], M.K. Prinsen [b], P.M. McNamee [c], R. Roggeband [d]

[a] Procter & Gamble, Product Safety and Regulatory Affairs, Temselaan 100, B-1853 Strombeek-Bever, Belgium
[b] TNO Quality of Life, Toxicology and Applied Pharmacology Department, Utrechtseweg 48, 3704 HE Zeist, The Netherlands
[c] The Procter & Gamble Company, Whitehall Lane, Egham, Surrey TW20 9NW, UK
[d] Procter & Gamble Service GmbH, Berliner Allee 65, 64274 Darmstadt, Germany

ARTICLE INFO

Article history:
Received 6 November 2008
Available online 19 May 2009

Keywords:
Isolated chicken eye test
ICE
Chicken enucleated eye test
CEET
Eye irritation
Household cleaning products
In vitro test methods
Alternative testing method
Low volume eye test
LVET

ABSTRACT

Eye irritation is an important endpoint in the safety evaluation of consumer products and their ingredients. Several *in vitro* methods have been developed and are used by different industry sectors to assess eye irritation. One such *in vitro* method in use for some time already is the isolated chicken eye test (ICE). This investigation focuses on assessing the ICE as a method to determine the eye irritation potential of household cleaning products, both for product safety assurance prior to marketing and for classification and labeling decisions. The ICE involves a single application of test substances onto the cornea of isolated chicken eyes. Endpoints are corneal swelling, corneal opacity and fluorescein retention. The ICE results were compared to historic LVET data in this study due to availability of such *in vivo* data and the ability to correlate LVET to human experience data on the outcome of accidental exposures to household cleaning products in general. The results of this study indicate that the ICE test is a useful *in vitro* method for evaluating the eye irritation/corrosion potential and establishing classification and labeling for household cleaning products. For new product formulations, it is best used as part of a weight-of-evidence approach and benchmarked against data from comparable formulations with known eye irritation/corrosion profiles and market experience.

1. Introduction

Historically the rabbit Draize eye irritation test has been used to assess eye irritation potential of substances and mixtures thereof. The assay is accepted by regulatory agencies worldwide (e.g., OECD, 2002; EC, 2004; USEPA, 1998) and is based on a method developed by Draize and colleagues in 1944 (Draize et al., 1944). The Draize test provides a quantitative scoring which is used as the basis for hazard classification of eye irritants and corrosives in international classification systems such as the European Union (EU) as well as under the United Nations (UN) Globally Harmonized System (GHS) classification and labeling scheme (EC, 2001; UNECE, 2003). Both classification systems are based on the severity of the ocular tissue lesions and/or persistence of effects. The EU hazard classification of eye irritants uses the risk phrases 'R36' (Irritating to eyes) and 'R41' (Risk of serious damage to eyes), based on whether the levels of damage, averaged across the 24, 48 and 72 h observation times for each ocular tissue lesion, fall within or above certain ranges of scores. The UN GHS considers two harmonized categories, one for irreversible effects/serious damage to the eye (Category 1), and one for reversible effects (Category 2). Reversible effects are further sub-classified, based on the duration of persistence (Category 2A: Irritating to eyes, reverses within 21 days and Category 2B: Mildly irritating to eyes, reverses within 7 days).

Though the Draize test has served the community well for decades there are, as with any assay, generally recognized limitations of this assay. Scientific publications describe challenges of the Draize test related to variability, subjectivity of scoring and over-prediction of the human response (Weil and Scala, 1971; York and Steiling, 1996; Buehler, 1974; Heywood and James, 1978; Jacobs et al., 1987; Daston and Freeberg, 1991). These challenges, added to concerns about animal welfare and a scientific desire to have available eye irritation assays that are based on better understanding of eye injury at the tissue and cellular level, have led researchers to investigate 3Rs alternative methods both *in vivo* (refinement) and *in vitro* (replacement) ones.

A number of *in vitro* methods, most notably organotypic models, have been evaluated for their ability to identify eye irritants/corrosives. Organotypic models employ eye tissues (e.g., isolated eyes and corneas) from food-chain animals and include the bovine corneal opacity and permeability test (BCOP), the isolated chicken eye

* Corresponding author. Fax: +32 2 568 35 91.
E-mail addresses: schutte.k@pg.com (K. Schutte), menk.prinsen@tno.nl (M.K. Prinsen), mcnamee.pm@pg.com (P.M. McNamee), roggeband.r@pg.com (R. Roggeband).

0273-2300/$ - see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.yrtph.2009.05.008

test (ICE), also known as chicken enucleated eye test (CEET) and the isolated rabbit eye test (IRE) all of which assess total corneal damage.

The ICE was introduced by Prinsen and Koëter in 1993 as a modification of the IRE (Burton et al., 1981). In brief, the ICE involves a single dose application of a test substance directly onto the cornea of isolated chicken eyes. The endpoints measured are corneal swelling, corneal opacity and fluorescein retention. Corneal swelling, measured as thickness, has been identified as a quantitative and reliable endpoint for the evaluation of corneal injury both *in vivo* and *in vitro* (Burton, 1972; Burton et al., 1981). Corneal opacity provides an assessment of corneal damage in the ICE that can be directly correlated to corneal damage observed in the *in vivo* rabbit eye test. Finally, fluorescein retention provides information on corneal permeability, indicative of damage to the corneal surface. The procedures for conduct of the ICE are described in INVITTOX protocol 80 (http://evcam-sis.jrc.it/invittox/published/indexed_80.html).

Early use of the ICE assay for 21 chemicals with a known Draize profile identified that the ICE test correctly classified all chemicals that require R36 or R41 classification within the EU (Prinsen and Koëter, 1993). Further, the assay showed good correlation with data obtained in Draize on industrial materials and certain formulations tested in standard contract toxicity evaluations at the TNO laboratories (Prinsen, 1996). Together with other alternative assays, the ICE has been reviewed in a range of validation or evaluation studies (Balls et al., 1999; Worth and Balls, 2002) with the outcome that no single test was found capable of fully replacing the Draize test, but some of the assays, including the ICE, showed considerable promise as screening tools for eye irritancy/corrosion. Most recently, the ICE along with other organotypic assays has been reviewed by the Interagency Coordination Committee on the Validation of Alternative Methods (ICCVAM, 2006). ICCVAM accepted the ICE in 2006 as a screening test to identify substances as ocular corrosives and severe irritants (i.e., EPA Category I, UN GHS Category 1, EU R41) in a tiered testing strategy and as part of a weight-of-evidence approach. The ECVAM Scientific Advisory Committee then endorsed this conclusion in 2007 (ECVAM, 2007). Furthermore, the ICE is now accepted by both EU and US regulatory authorities for this purpose.

This study examined the suitability of the ICE test for its potential to predict eye irritation/corrosion and its utility for classification and labeling in the context of household cleaning products. For this purpose, a total of 20 household cleaning products and raw materials were tested in the ICE test. The results were compared with *in vivo* data from a rabbit assay that is a refinement of the Draize test – the low volume eye test (LVET). It is recognized that LVET is not a regulatory approved *in vivo* eye irritation assay. However, for the purpose of this evaluation for household and cleaning products this was considered to be an appropriate *in vitro* to *in vivo* correlation for several reasons that include: (1) LVET is an established, mechanistically-based (Maurer et al., 2002) *in vivo* eye irritation test that uses biological and physiological endpoints that are relevant to humans for which there is an identified prediction model that is the same as that for the Draize test; (2) with the exception of the dosing regimen in which a lower volume (10 μL versus 100 μL) of test material is instilled directly onto the cornea instead of into the lower conjunctival sac, LVET is the same assay as the Draize test in terms of the ocular tissues evaluated (cornea, conjunctiva, iris), scoring system and interpretation of individual ocular tissues data used for calculation of a classification/labeling within different regulatory schemes; (3) considering anatomical and physiological differences between species (rabbit and human) the dose volume of 10 μL in LVET is an appropriate dose volume (Swanston, 1985; Mishima et al., 1966; Ehlers, 1976; Chrai et al., 1973); (4) LVET has been correlated as being predictive of the human response from accidental exposure to household and cleaning products through clinical studies (Freeberg et al., 1986a; Ghassemi et al., 1993; Roggeband et al., 2000) and human experience (accidental consumer and industrial exposures and Poison Control Centres) (Freeberg et al., 1984, 1986b; Cormier et al., 1995) and (5) the LVET has been used successfully for many years by members of the household and cleaning products industry to support the consumer safety of such products. Since the ultimate objective here is to predict the human eye response, the ICE data have been compared with historically available LVET data.

2. Materials and methods

2.1. Test substances

The study was conducted with 15 common household cleaning products and five raw materials, for most of which historical LVET data was available. The test samples varied in terms of their formulation characteristics (i.e., low and high pH, bleaches, surfactant-based liquids or powders) and their potential to cause eye irritation as observed in the *in vivo* LVET assay. The principal test sample characteristics as well as the availability of LVET results or literature data are presented in Table 1.

2.2. ICE procedure

Approximately 7 week old male or female ROSS spring chickens of 2.5–3.0 kg bodyweight from the slaughterhouse were used as eye donors. Chicken heads were taken immediately after sedation of the animals by electric shock and incision of the neck for bleeding and transported to the test facilities of TNO Quality of Life, Toxicology and Applied Pharmacology, Zeist, The Netherlands. Within 2 h after kill, the eyes were dissected and placed in a superfusion apparatus as described below.

In a first step, the eyelids were carefully removed from the chicken head without damaging the cornea and a small drop of fluorescein sodium BP 2% w/v (Minims™, disposable single-use droppers, Smith & Nephews Ltd., Romford, England) was applied to the corneal surface for a few seconds, then rinsed off with isotonic saline at ambient temperature. The fluorescein-treated cornea was examined with a slit-lamp microscope (Slit-lamp 900 CN, Haag-Streit AG, Liebefeld-Bern, Switzerland) to ensure that no damage occurred. Undamaged eyes were then carefully removed from the head, placed in a stainless steel clamp with the cornea positioned vertically and transferred to one of the eleven chambers of the superfusion apparatus. The clamp was positioned in such a way that the entire cornea was supplied with an isotonic saline drop at a rate of ca. 0.10–0.15 mL/min through a peristaltic pump (Watson-Marlow 295CA, Rotterdam, The Netherlands). The chambers of the superfusion apparatus and the saline were kept at 32 ± 1.5 °C (water pump, Thermomix 1441, B. Braun Melsungen AG, Melsungen, Germany).

The studies were conducted with three test eyes per sample and one control eye per assay (one assay covered three samples). One control eye was used to demonstrate the suitability of the general conditions in the superfusion apparatus during the test period, i.e., saline drip and temperature. Additional control eyes were not required because each test eye acted as its own control, by providing baseline values for corneal swelling, corneal opacity and fluorescein retention prior to dosing. Before test start, all eyes were examined once again with the slit-lamp to ensure they were not damaged. Corneal thickness was measured at the corneal apex using the Depth Measuring Attachment No. I for the Haag-Streit slit-lamp microscope. Eyes with a corneal thickness deviating more than 10% from the mean value, unacceptably stained with fluorescein (score higher than 0.5, indicating a permeable cornea) or showing other signs of damage were discarded.

After an equilibration period of 45–60 min, corneal thickness was measured again to determine the zero reference value for corneal swelling calculations. Immediately afterward, at time $t = 0$, the

Table 1
Test samples and availability of LVET results.

Test sample	Product category	LVET *in vivo* data available?
Household cleaning products		
Liquids		
Acidic cleaner (1)	Low pH liquid (pH = 1.7 at 100%) (acid, stabilizers, perfume)	Yes
Acidic cleaner (2)	Low pH liquid (pH = 3.0 at 100%) (acid, stabilizers, perfume)	Yes
Fabric softener	Low pH liquid (pH = 3.5 at 3%) (cationic surfactant-based, stabilizers, perfume)	Yes[a]
Acidic peroxide bleach	Low pH liquid (pH = 4.0 at 100%) (hydrogen peroxide-based, surfactant, perfume)	Yes
Alkaline cleaner	High pH liquid, aqueous (solvent, alkali, stabilizers, surfactant, perfume) (pH = 12 at 10%)	Yes
Alkaline bleach cleaner	High pH liquid, (hypochlorite-based, surfactants) (pH 11 at 1%)	Yes
Automatic dishwashing bleach gel	High pH dishwasher gel (pH = 11.7 at 1%) (hypochlorite, silicates, alkali, stabilizers)	Yes
Dishwashing liquid	Surfactant-based liquid (nonionic and anionic surfactants, stabilizers, perfume) (pH = 9 at 10%)	Yes
Powders		
Powder detergent (1)	Laundry detergent powder (surfactants, builders, chelants, polymers, perfume) (pH = 10.5 at 1%)	Yes
Powder detergent (2)	Detergent powder like (1) plus 1% sulphamic acid	No[b]
Powder detergent (3)	Detergent powder like (1) plus 5% sulphamic acid	No[b]
Powder detergent (4)	Detergent powder like (1) plus 7% sulphamic acid	No[b]
Powder detergent (5)	Detergent powder like (1) plus 5% citric acid	No[b]
Automatic dishwashing powder	High pH powder based on silicates, alkali, stabilizer and surfactants (pH = 10.7 at 1%)	Yes
Bleach additive powder	Percarbonate bleach-based laundry additive powder (pH = 10 at 1%)	Yes
Raw materials		
Powders		
Bleach catalyst	Powder raw material, confidential (pH = 5.2 at 1%)	Yes
Citric acid	Powder raw material, pure substance (pH = 2.5 at 1%)	No[c]
Sulphamic acid	Powder raw material, pure substance (pH = 1.2% at 1%)	No[c]
Silicate 2-ratio	Powder raw material, pure substance (pH = 11.7 at 1%)	Yes
Metasilicate 1-ratio	Powder raw material, pure substance (pH = 12.8 at 1%)	Yes

[a] LVET data available on a closely related formulation.
[b] LVET data on powder detergent (1) are used as reference.
[c] Literature data (Draize) are used as reference.

test substance was applied. For this, the clamp holding the eye was placed on tissue paper outside the chamber with the cornea facing upwards. The standard testing protocol involves the application of either 30 μL of a liquid test substance or 30 mg of a solid test substance to cover the entire surface of the cornea. This volume of 30 μL was selected for the standard ICE protocol to mimic the Draize test where 100 μL are used, taking into account that the chicken cornea is approximately 1/3 the size of a rabbit cornea. In this study, additional volumes (i.e., 3 and/or 10 μL) or masses (i.e., 3 and/or 10 mg) of the test samples were also applied, in order to evaluate if they would provide better comparisons to the LVET assay (where 10 μL are used).

Ten seconds after application, the corneal surface was thoroughly rinsed by application of exactly 20 mL of isotonic saline and the eyes were then returned to the superfusion chamber. Using a slit-lamp microscope, corneal swelling and corneal opacity were determined after 30, 75, 120, 180 and 240 min, and fluorescein retention after 30 min.

At the end of the study, test and control eyes were preserved in a neutral aqueous phosphate-buffered 6% formaldehyde solution, to be later embedded in paraffin wax, sectioned at 5 μM and examined histologically for morphological effects after staining with hematoxylin and eosin.

Table 2
Determination of the category scores for corneal swelling, corneal opacity and fluorescein retention.

Corneal swelling (Max. mean % swelling)	Corneal opacity (Max. mean opacity score)	Fluorescein retention (Mean retention score)	Corresponding category
0–5	0.0–0.5	0.0–0.5	Category I (no effect)
>5–12	0.6–1.5	0.6–1.5	Category II (slight effect)
>12–18[a]			
>12–18[ba]	1.6–2.5	1.6–2.5	Category III (moderate effect)
>18–26			
>26–32			
>26–32[b]	2.6–4.0[c]	2.6–3.0	Category IV (severe effect)
>32			

Prediction scheme is described in Prinsen and Koëter (1993) and Prinsen (1996).
[a] >75 min after treatment.
[b] <75 min after treatment.
[c] In case of score 4, thickness assessment not possible.

2.3. Criteria and scoring system

The severity level for each study endpoint was evaluated for the three test eyes of each sample according to the following set of criteria and scoring systems.

2.3.1. Corneal swelling

The mean percentage of corneal swelling was calculated for each observation time point as follows:

$$\frac{\text{corneal thickness at time } t - \text{corneal thickness at time} = 0}{\text{corneal thickness at time} = 0} \times 100$$

Based on the highest mean value obtained at any of the observation time points, a category score for corneal swelling was then determined, as shown in Table 2.

2.3.2. Corneal opacity

Corneal opacity was defined as 'opacity degree of density' and assessed by scoring the area of the cornea that was most densely opacified.

Score	Observation
0	No opacity
0.5	Very faint opacity
1	Scattered or diffuse areas; details of the iris are clearly visible
2	Easily discernible translucent area; details of the iris are slightly obscured
3	Severe corneal opacity; no specific details of the iris are visible; size of the pupil is barely discernible
4	Complete corneal opacity; iris invisible

Mean corneal opacity was calculated for each observation time point. Based on the highest mean score obtained at any of the observation time points, a category score for corneal opacity was determined (Table 2).

2.3.3. Fluorescein retention

Fluorescein retention was scored as shown below:

Score	Observation
0	No fluorescein retention
0.5	Very minor single cell staining
1	Single cell staining scattered throughout the treated area of the cornea
2	Focal or confluent dense single cell staining
3	Confluent large areas of the cornea retaining fluorescein

When test substances adhered to the cornea, fluorescein retention was determined after adequate removal of the test substance. Based on the mean fluorescein retention score obtained at 30 min, a category score for corneal opacity was determined (Table 2).

2.3.4. Morphological effects

Morphological changes in the test eyes were recorded for each of the test substances. The effects included pitting of corneal epithelial cells, loosening of the epithelium, roughening of the corneal surface and sticking of the test substance to the cornea. These findings could vary in severity and could occur simultaneously.

2.3.5. Microscopic effects

Corneal lesions were determined by histological examination. The effects included but were not limited to erosion, necrosis and vacuolation of the epithelium, disorder of stromal fibers, pycnotic nuclei in the stroma (anterior/posterior region) and necrosis of the endothelium. The description of these findings was subjective to the interpretation of the investigator.

2.3.6. ICE overall eye irritancy categorization

Based on the category scores obtained for corneal swelling, corneal opacity and fluorescein retention and, if present, on morphological effects, an overall eye irritancy class was established for each of the products tested. As shown in Table 3, the substances were classed as not irritating, slightly, moderately or severely irritating, depending on the outcome of evaluation of the ICE endpoints.

2.3.7. Comparison of ICE and EU/UN GHS eye irritation classification results

In order to assess the suitability of the ICE test for determining eye irritation classification, the results obtained in this study were translated into the corresponding EU classification based on a conversion scheme developed using scientific judgment and many years of experience (Prinsen and Koëter, 1993; Prinsen, 1996, 2004). In addition, the general eye irritancy categorization scheme of the ICE also allowed for translation to UN GHS classification. The conversion scheme is shown in Table 3.

2.4. LVET procedure

The LVET procedure was conducted in accordance with the standard protocol published by the American Society for Testing and Materials originally in 1985 and reapproved in 2003 (ASTM, 2003). Three animals were used in each LVET. A preliminary macroscopic examination of the eyes of each rabbit was conducted using fluorescein dye. A minimum of 1 h after the preliminary ocular examination, the test article was placed directly on the cornea of the right eye of each animal. Liquids were administered at a volume of 10 μL using a glass syringe. Solids were administered as a weight equivalent to 10 μL volume not to exceed 10 mg. Following instillation, the eyelids were released without forced blinking or manipulation. The contra-lateral eye remained untreated to serve as a control.

Responses of the cornea, iris and conjunctiva in the test and control eyes of each rabbit were evaluated macroscopically using an auxiliary light source at 1, 24, 48 and 72 h after dosing according to the Draize scale for scoring ocular lesions (17). Following macroscopic observations at the 24-h scoring interval, the fluorescein examination procedure was repeated on all test and control eyes and any residual test article gently rinsed from the eye at this time (if possible) using physiological saline. If any fluorescein findings were noted, a fluorescein examination was further conducted at each subsequent interval until a negative response was obtained and/or until all corneal opacity had cleared. If there was no evidence of treatment-related ocular irritation at the 72-h scoring interval, the study was terminated. If ocular irritation persisted in any test eye, the observation period was extended for the affected animals (scored on days 7, 10, 14 and 21). Animals requiring an extended observation period remained on test until the irritation had resolved or permanent injury was evident.

2.4.1. Scoring system

The Draize scale for scoring ocular lesions was used to evaluate the effect on the ocular tissues (cornea, conjunctiva and iris) of exposure to test material (Draize, 1959). The scoring scale used here was the same as that which is used in the Dangerous Substances Directive Annex V Draize test (OECD, 2002).

2.4.2. Data interpretation

2.4.2.1. Ocular evaluation. Using the Draize scoring scale, the group mean irritation score was calculated for each scoring interval based on the number of animals initially dosed in each group. The calculated group mean ocular irritation scores for each interval was then used to categorise the test article according to the ocular evaluation criteria as defined by Kay and Calandra (1962).

2.4.2.2. EU ocular evaluation. The total ocular irritation score for the 24, 48 and 72 h intervals were individually added for corneal opacity, iris lesion, conjunctival redness and conjunctival edema. For a test containing three rabbits, the group mean scores for corneal opacity, iris lesion, conjunctival redness and conjunctival swelling were then calculated. The resulting mean ocular irritation scores were then classified according to the existing European Union (EU) hazard classification and labeling scheme within chemicals legislation (European Union, 2001).

3. Results

The results obtained in the ICE protocol for all 20 test materials are summarized in the Tables 4–6. Table 4 presents the highest mean scores for corneal swelling and corneal opacity, and the mean score for fluorescein retention, as well as the ICE eye irritation categorization derived therefrom. The acidic to neutral test materials generally caused mostly slight to moderate irritation effects, while the clearly alkaline cleaner formulations as well as powder raw material sulphamic acid and metasilicate produced severe irritation effects. Related histopathological findings for all test materials (only for eyes treated with 30 μL or mg) are given in Table 5. Finally, Table 6 then compares regulatory classifications derived from the ICE test results and those derived from LVET data previously. This comparison is visually illustrated in Fig. 1.

A direct comparison of the ICE and LVET-based EU classifications was possible for 14 out of the 20 products tested. For 4 products (powder detergents (2)–(5)), no LVET testing was conducted, therefore LVET data from a historic and similar formulation, powder detergent (1), was used as the *in vivo* reference. For the two acids, literature data on *in vivo* studies were used as reference.

Overall, as summarized in Table 6, the ICE results were either in line with or more conservative than the eye irritation profile predicted on the basis of the LVET test. When a dosing volume of 30 μL or mg was applied, 14 out of 20 test samples revealed ICE-based regulatory classifications which were comparable to those of the LVET, 5 out of 20 were over-predicted and only one was under-predicted (silicate 2-ratio powder). At 10 μL or mg, the ICE-based regulatory classification was in line with LVET in 13 out of 20 cases, over-predicted in 4 out of 20 cases, and under-predicted 3 out of 20 times. These data are illustrated in a visual way in Fig. 1, where those test substances which showed a match between ICE and LVET in terms of irritation category results are highlighted in green. At 3 μL or mg there were only few data points; results were generally identical to what was seen at 10 μL or mg. As a general trend, there was a good match between ICE and LVET results for surfactant-based and low pH products. High pH products showed more often over-prediction in the ICE versus the LVET assay.

At the histological level, all but two liquids ('acidic peroxide bleach' and 'dishwashing liquid'), showed some degree of effect on the epithelium, mainly necrosis and vacuolation. A few samples also induced pycnotic nuclei in the outer region of the stroma. For the two remaining liquids mentioned above no effects were seen. All powder products had very slight to moderate effects on the epithelium. Several powders also led to pycnotic nuclei in the anterior region of the stroma. Metasilicate 1-ratio was the only test substance to induce effects in the posterior region of the stroma and the endothelium.

4. Discussion

The ICE test was accepted in a range of EU Member States to identify substances as ocular corrosives and severe irritants in a tiered testing strategy and as part of a weight-of-evidence approach, even before and also in line with the ICCVAM (2006) and ECVAM (2007) conclusions on the assay. It has been used by industry for screening purposes, predominantly early in the product development cycle to identify product formulations that may have an unacceptable eye irritation/corrosion profile. Whereas parallel *in vitro* ICE and *in vivo* Draize eye irritation testing (OECD 405) showed good correlation with about 100 test substances (Prinsen, 1996), it has proven difficult to compare ICE test results with those of *in vivo* eye irritation tests of various other sources (ICCVAM, 2006; Worth and Balls, 2002). This could in part be due to reported limitations of the *in vivo* tests (e.g., the Draize test), particularly (1) the often unknown exposure conditions (up to 24 h for solids), and (2) the subjective scoring of tissue lesions in the Draize test resulting in variable estimates of eye irritancy (Prinsen, 2006).

This study examined the suitability of the ICE test for its potential to predict eye irritation/corrosion and its utility for

Table 3
Overall ICE eye irritancy categorization based on the scores for corneal swelling, corneal opacity and fluorescein retention, and consequent conversion into corresponding EU and UN GHS classification.

Combination of category scores	ICE general eye irritancy category	Corresponding UN GHS eye irritancy classification	Corresponding EU eye irritancy classification
2 x I, 1 x II	Not irritating	Not irritating	Not irritating
3 x I	Not irritating	Not irritating	Not irritating
3 x II[b]	Slightly irritating	2B	Not irritating
2 x II, 1 x I	Slightly irritating	2B	Not irritating
1 x I, 1 x II, 1 x III[a]	Slightly irritating	2B	Not irritating
2 x II, 1 x III[c]	Slightly irritating	2B	R36
2 x I, 1 x IV[a]	Slightly irritating	2B	R36
3 x III	Moderately irritating	2A	R36
2 x III, 1 x II	Moderately irritating	2A	R36
2 x III, 1 x IV[d]	Moderately irritating	2A	R36
2 x III, 1 x I[a]	Moderately irritating	2A	R36
2 x II, 1 x IV[a]	Moderately irritating	2A	R36
1 x II, 1 x III, 1 x IV[a]	Moderately irritating	2A	R36
3 x IV	Severely irritating	1	R41
2 x IV, 1 x III	Severely irritating	1	R41
2 x IV, 1 x II[a]	Severely irritating	1	R41
2 x IV, 1 x I[a]	Severely irritating	1	R41
Immediate(after application) corneal opacity score of 3	Severely irritating	1	R41
Corneal opacity score of 4 (at any time)	Severely irritating	1	R41
Severe loosening of epithelium			

Prediction scheme is described in Prinsen and Koëter (1993) and Prinsen (1996).
[a]Combinations of these categories are less likely to occur.
[b]Combination can be considered a borderline case between non-irritating and irritating (R36).
[c]Combination can be considered as a borderline case between mildly irritating (2B) and irritating (2A).
[d]Combination can be considered a borderline case between irritating (R36/2A) and severely irritating (R41/1).

Table 4
ICE test results – scores and eye irritation categorization.

Test sample	Mean scores											
	Corneal swelling			Corneal opacity			Fluorescein retention			Eye Irritancy category based on ICE		
Liquids (μL)	**3**	**10**	**30**	**3**	**10**	**30**	**3**	**10**	**30**	**3**	**10**	**30**
Acidic cleaner (1)	7	13	21	1	1	2	2	2	2.3	Slight	Slight	Moderate
Acidic cleaner (2)	-	12	15	-	1.2	2.5	-	2	2	-	Slight	Moderate
Fabric softener	-	5	8	-	1	1	-	1	1	-	Slight	Slight
Acidic perox. bleach	-	17	22	-	0.5	1	-	1	1	-	Slight	Slight
Alkaline cleaner	38	40	44	3	3	3	3	3	3	Severe	Severe	Severe
Alkaline bleach cleaner	-	24	25	-	2.8	3.0	-	2.5	2.5	-	Moderate	Moderate
Automatic dish-washing bleach gel	-	26	49	-	4	4	-	3	3	-	Severe	Severe
Dishwashing liquid	-	7	23	-	0.5	1.2	-	0.7	1.3	-	Slight	Slight
Powders (mg)	**3**	**10**	**30**	**3**	**10**	**30**	**3**	**10**	**30**	**3**	**10**	**30**
Powder detergent (1)	3	8	12	0.5	1	1	1	0.8	1	None	Slight	Slight
Powder detergent (2)	-	2	6	-	1	1.5	-	1.3	1	-	Slight	Slight
Powder detergent (3)	-	5	12	-	0.8	1.5	-	1	0.8	-	Slight	Slight
Powder detergent (4)	-	6	7	-	1	1	-	1.3	1.3	-	Slight	Slight
Powder detergent (5)	-	6	9	-	1.2	1.0	-	1.3	1.3	-	Slight	Slight
Automatic dish-washing powder	3	5	3	1	1	1	1	1	1	Slight	Slight	Slight
Bleach additive powder	6	6	13	1	1	1.3	1.5	1.5	1.5	Slight	Slight	Slight
Bleach catalyst	-	2	3	-	0.5	0.3	-	0.5	0.5	-	None	None
Citric acid	-	6	9	-	1.8	2	-	2	1.7	-	Moderate	Moderate
Sulphamic acid	-	8	11	-	3	3	-	2	2	-	Severe[a]	Severe[a]
Silicate 2-ratio	1	1	1	0.5	1	1	0.7	0.8	1.0	-	Slight	Slight
Metasilicate 1-ratio	54	53	53	4	4	4	3	3	3	Severe	Severe	Severe

[a]Based on immediate opacity (score 3) after application.

classification and labeling in the context of household cleaning products which were previously mostly evaluated in the low volume eye test (LVET) and for which good quality data were available. The LVET is a modification of the Draize protocol and uses a lower dose volume of 10 μL (versus 100 μL in the Draize) that better fits the anatomical and physiological characteristics of the human eye. In addition to giving responses closer to those observed in humans, it has been used to provide data on mechanistic understanding of the eye response to chemical injury (Maurer et al., 2002; Jester, 2006). The LVET was developed over 25 years ago by Griffith et al. (1980). Since then, it has been correlated as being predictive of the human response from accidental exposure to household and cleaning products through clinical studies (Roggeband et al., 2000) and through decades of market place experience of the Procter & Gamble Company and other members of industry.

For a total of 20 household cleaning products and raw materials, a comparison was made between the EU and UN GHS classifications derived using ICE and LVET test data (see Table 6). The outcome of this investigation shows that the eye irritancy/corrosion classifications obtained for the sampled products based on ICE results were, in a majority of cases, either in line with or more conservative than the eye irritation profile predicted on the basis of the LVET test. Some conservatism or over-prediction was in fact to be expected because the ICE model was developed to predict the Draize test. The dose of 30 μL or mg proved to yield better and more consistent results than 10 or 3 μL or mg. As a general trend, there was a good match between ICE and LVET results for surfactant-based and low pH products, both in liquid as well as in powder form. High pH products showed more often over-prediction in the ICE as compared to the LVET assay, especially the three liquid ones alkaline cleaner, alkaline bleach cleaner and automatic dishwashing bleach gel. Only one sample, i.e., silicate 2-ratio, was under-predicted at the 30 mg dose; this can be explained by the fact that this raw material is a powder that dissolves only slowly in water and therefore did not lead to irritating effects during the standard exposure time of 10 s. It was however correctly predicted upon longer exposure (data not shown).

This study confirms the suitability of the ICE test for the determination of the eye irritation/corrosion potential of household cleaning products and raw materials when compared with LVET which is the *in vivo* assay for which most data are available for household and cleaning products and which has been successfully correlated with the human response through both clinical studies and many years of human experience for these types of products. In the frame of product developments, ICE could be an appropriate assay to determine the profile of new cleaning product formulations by comparison to similar formulations for which historic *in vivo* eye irritation data and market experience exists. ICE data are regularly used by Procter & Gamble in a weight-of-evidence approach to determine the eye safety and appropriate classification of new candidate cleaner and detergent formulations. The ICE test is also useful to screen candidate formulations. Where for example several prototypes containing various concentrations of the same ingredients are being developed, ICE test results can help select the prototype(s) with the best irritation versus efficacy profile. Recent examples for which the ICE test was used within a weight-of-evidence approach for risk assessment and classification decisions included an acidic cleaner, a powder detergent and a concentrated liquid detergent. These three products were related to the test materials evaluated in the study but explored new innovation directions.

This study further suggests the utility of the ICE test for eye irritancy/corrosion classification and labeling purposes in the frame of existing classification schemes (e.g., EU, UN GHS). ICE results can be translated into the EU and UN GHS eye irritation schemes (as per Table 3) and benchmarked against data for

Table 5
Morphological findings for chicken eyes treated with 30 μL or mg of test substance (figures in brackets are number of eyes affected out of three; only indicated if different from 3/3).

Test Sample	Epithelium			Stroma			Endothelium
	Erosion	Necrosis	Vacuolation	Disorder of fibres	Pycnotic nuclei anterior region	Pycnotic nuclei posterior region	Necrosis
Liquids							
Acidic cleaner (1)	-	½ (2/3) 1 (1/3)	2	-	-	-	-
Acidic cleaner (2)	-	½ (1/3) 1 (1/3) 2 (1/3)	1 (1/3) 2 (1/3) 3 (1/3)	-	½ 1 (1/3)	-	-
Fabric softener	½	½	½ (2/3) 1 (1/3)	-	-	-	-
Acidic peroxide bleach	-	-	-	-	-	-	-
Alkaline cleaner	-	-	- (1/3) 1 (1/3) 2 (1/3)	-	1 (1/3) 2 (2/3)	-	-
Alkaline bleach cleaner	- (1/3) ½ (2/3)	1 (1/3) 2 (2/3)	2	-	1	-	-
Automatic dish-washing bleach gel	1	-	1 (2/3) 2 (1/3)	-	-	-	-
Dishwashing liquid	-	-	-	-	-	-	-
Powders							
Powder detergent (1)	½	- (1/3) 1 (2/3)	-	-	- (2/3) 1 (1/3)	-	-
Powder detergent (2)	½	½	½	-	½ (2/3) 1 (1/3)	-	-
Powder detergent (3)	½	-	-	-	- (2/3) ½ (1/3)	-	-
Powder detergent (4)	½	- (2/3) ½ (1/3)	- (2/3) ½ (1/3)	-	- (2/3) ½ (1/3)	-	-
Powder detergent (5)	½	½	-	-	-	-	-
Automatic dish-washing powder	½	½	2	-	-	½	½
Bleach additive powder	½	-	1	-	-	-	-
Bleach catalyst	½ (2/3) 1 (1/3)	-	-	-	- (2/3) ½ (1/3)	-	-
Citric acid	1	-	1 (2/3) 2 (1/3)	-	-	-	-
Sulphamic acid	½	½	½ (2/3) 1 (1/3)	-	-	-	-
Silicate 2-ratio	½	- (1/3) ½ (2/3)	-	-	-	-	-
Metasilicate 1-ratio	1 2	1 (1/3) 2 (2/3)	No data	-	1 (1/3) 2 (2/3)	½ (1/3) 1 (1/3) 2 (1/3)	2 (1/3) 3 (2/3)

– = not observed; ½, very slight; 1 = slight; 2 = moderate; 3 = severe.

comparable formulations for which EU classifications have been derived.

Furthermore, microscopic evaluation of the cornea can provide important information on the severity of the injury at the tissue and cellular level. Maurer et al. have proposed that differences in ocular irritation are related to differences in extent of initial injury, and that regardless of the process leading to tissue damage, extent of injury is a principal factor determining the outcome of ocular irritation (Maurer et al., 2002, which reviews all studies conducted in the mechanistic program). Using light microscopy they characterized the general ocular irritancy of a broad sampling of surfactants and non-surfactants (acid, alkali, ketone, alcohol, aromatic amine, aldehyde) causing slight to severe irritation in the context of the LVET. Additionally, using *in vivo* confocal microscopy, they quantitatively characterized the corneal changes occurring with these irritants in the LVET. Despite differences in the mechanisms by which these materials may cause injury, the results collectively support their hypothesis that by defining the initial extent of injury associated with ocular irritation it is possible to predict the subsequent response and final outcome. Importantly, they proposed that

Table 6
Comparison of EU and GHS classification based on ICE versus LVET test results.

Test sample	Characteristics	CCM	EU/GHS classification based on ICE test results			EU/GHS classification based on LVET test results	
Liquids (µl)			3	10	30	EU	GHS
Acidic cleaner (1)	Low pH	R36	R36/2B	R36/2B	R36/2B	NC [a]	NC
Acidic cleaner (2)	Low pH	NC	- [b]	R36/2B	R36/2A	NC	NC
Fabric softener	Low pH	NC	- [b]	NC/2B	NC[1]/2B	NC	NC
Acidic peroxide bleach	Low pH	R41	- [b]	NC/2B	R36/2B	R36	2A
Alkaline cleaner	High pH	R41	R41/1	R41/1	R41/1	R36	2B
Alkaline bleach cleaner	High pH	R41	- [b]	R36/2A [c]	R36/2A [c]	NC	2B
Automatic dishwashing bleach gel	High pH	R41	- [b]	R41/1	R41/1	NC	2A
Dishwashing liquid	Surfact.-based	R41	- [b]	NC/2B	R36/2B	R36/41	2A/1
Powders (mg)			3	10	30		
Powder detergent (1)	Surfact.-based	R41	NC	NC[a]/2B	NC[1]/2B	NC	NC
Powder detergent (2)	Surfact.-based	R41	- [b]	NC/2B	NC[1]/2B	NC [d]	NC [d]
Powder detergent (3)	Surfact.-based	R41	- [b]	NC/2B	NC[1]/2B	NC [d]	NC [d]
Powder detergent (4)	Surfact.-based	R41	- [b]	NC[1]/2B	NC[1]/2B	NC [d]	NC [d]
Powder detergent (5)	Surfact.-based	R41	- [b]	NC[1]/2B	NC[1]/2B	NC [d]	NC [d]
Automatic dish-washing powder	High pH	R41	NC/2B	NC/2B	NC/2B	NC	NC
Bleach additive powder	High pH	R41	NC/2B	NC/2B	NC[1]/2B	NC	NC
Bleach catalyst	High pH	NA	- [b]	NC/NC	NC/NC	NC	2B
Citric acid	Low pH	NA	- [b]	R36/2B	R36/2A	R36 [e]	2B [e]
Sulphamic acid	Low pH	NA	- [b]	R36/2A [c]	R36/2A [c]	R36/41 [e]	2A/1 [e]
Silicate 2-ratio	High pH	NA	NC	NC[a]/2B	NC[a]/2B	R36	2A
Metasilicate 1-ratio	High pH	NA	R41/1	R41/1	R41/1	R41	1

CCM – conventional classification method according to Directive 1999/45/EC: The CCM is a method described in the EU Dangerous Preparations Directive (1999/45/EC) to estimate the theoretical hazard of a preparation based on a simple summation of the individual hazards of the substances in that preparation. NC, Not classified; NA, Not applicable as CCM method cannot be applied to substances.
[a] NC but close to R36.
[b] Dose volume not tested.
[c] R36/2A but close to R41/1.
[d] Powder detergent (1) used as LVET reference.
[e] Estimated based on literature data. Green boxes indicate cases where the ICE and LVET results are in line. Yellow and orange are when the ICE over or underpredicts, respectively, compared to the LVET.

a mechanistically based approach to the development of alternative ocular irritation tests would be to include histological evaluation to dimension the extent of initial injury using either *ex vivo* or *in vitro* corneal equivalent systems.

In the context of eye safety risk assessment or classification decisions, such histological assessments could add to the weight-of-evidence analysis in assays such as the ICE, especially when decisions between non-irritant and irritant (i.e., not classified or R36/2A) and between moderate and severe irritants (i.e., R36/2A or R41/1) have to be made. In borderline cases, a decision towards a more severe eye irritant categorization or classification is justified the more pronounced and deeper the injury. This is based on mechanistic work conducted by Maurer et al. suggesting that the deeper the corneal injury, the smaller the likelihood of fast and complete recovery (Maurer et al., 2002; Jester, 2006). For example, pronounced damage to the epithelium and the posterior region of the stroma or at endothelial level indicate high irritation potential justifying a more severe eye irritant categorization or a R41/1 classification. The mechanistic work also indicates that damage to the epithelium and anterior stromal region alone is typical of mild–moderate irritants and linked with potential for good recovery.

5. Conclusions

This study indicates that the ICE test is a useful *in vitro* methodology for evaluating the eye irritation/corrosion potential and establishing the EU classification and labeling of powder and liquid household cleaning products when compared with LVET which is the *in vivo* assay for which most data are available for household and cleaning products and which has been successfully correlated with the human response through both clinical studies and many years of human experience for these types of products. The good correlation of ICE results with LVET data for an identified range of products (in particular low pH and surfactant-based products) suggest that the ICE can be used to predict eye irritation responses in man for these product types. For new product formulations, it is best used as part of a weight-of-evidence tiered approach and benchmarked against available data from comparable formulations with known eye irritancy/corrosion profile and market experience. The number of samples comprised in the present research program was limited; further investigations will be conducted to better qualify the ICE test for this purpose and evaluate whether its use can be extended to a wider range of product types.

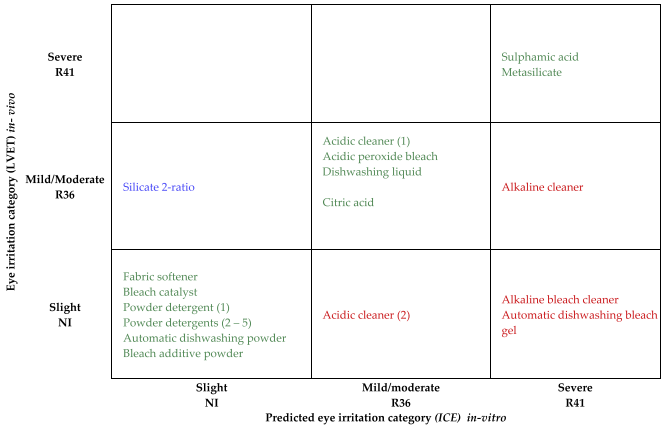


Fig. 1. Visualization of the data reported in Table 6 (green: matching prediction, blue: underprediction, red: overprediction).

References

American Society for Testing and Materials, 2003. Standard Method for Evaluation of Eye Irritation in Albino Rabbits, E1055-93 (reapproved in 2003) in Annual Book of ASTM Standards. Philadelphia, PA.

Balls, M., Berg, N., Bruner, L.H., Curren, R.D., de Silva, O., Earl, L.K., Esdaile, D.J., Fentem, J.H., Liebsch, M., Ohno, Y., Prinsen, M.K., Spielmann, H., Worth, A.P., 1999. Eye irritation testing: the way forward. The report and recommendations of ECVAM workshop 34. Altern. Lab. Anim. 27, 53–77.

Buehler, E.V., 1974. Testing to predict potential ocular hazards of household chemicals. In: Winek, C.L. (Ed.), Toxicology Annual 1974. Marcel Dekker, p. 53.

Burton, A.B.G., 1972. A method for the objective assessment of eye irritation. Food Cosmet. Toxicol. 10, 209.

Burton, A.B.G., York, M., Lawrence, R.S., 1981. The *in vitro* assessment of severe eye irritants. Food Cosmet. Toxicol. 19, 471–480.

Cormier, E.M., Hunter, J.E., Billhimer, W., May, J., Farage, M., 1995. Use of clinical and consumer eye irritation data to evaluate the Low Volume Eye Test. J. Toxicol. Cut. Ocular Toxicol. 14, 197–205.

Chrai, S., Patton, T.F., Mehta, A., Robinson, J.R., 1973. Lacrimal and instilled fluid dynamics in rabbit eyes. J. Pharmaceut. Sci. 62, 1112–1121.

Daston, G.P., Freeberg, F.E., 1991. Chapter 16: ocular irritation testing. In: Hobson, D.W. (Ed.), Dermal and Ocular Toxicology. Fundamentals and Methods. CRC Press, Boca Raton, pp. 510–539.

Draize, J.H., Woodward, G., Calvery, H.O., 1944. Methods for the study of irritation and toxicity of substances applied directly on the skin and mucous membranes. J. Pharmacol. Exper. Ther. 82, 377–390.

Draize, J.H., 1959. Appraisal of the Safety of Chemicals in Food, Drugs and Cosmetics. Association of Food and Drug Officials of the US Editorial Committee, 1959. Second Printing 1965. Topeka, Kansas. pp. 46–51.

Ehlers, N., 1976. Pharmacology of the conjunctival sac. In: Dickstein, S. (Ed.), Drugs and Ocular Tissues (Second Meeting of the International Society for Eye Research, Jerusalem, 1976). Karger, S.Basel. p. 23.

European Center for the Validation of Alternative Methods (ECVAM), 2007. ESAC statement on the conclusions of the ICCVAM retrospective study on organotypic in vitro assays as screening tools to identify potential ocular corrosives and severe irritants as determined by USEPA, EU (R41) and UN GHS classifications in a tiered testing strategy, as part of a weight of evidence approach. <http://ecvam.jrc.it/publication/ESAC26_statement_Organotypic_20070510c.pdf>(accessed July 2008).

European Commission (EC), 2001. Commission Directive 2001/59/EC of 6 August 2001 adapting to technical progress for the 28th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Communities L255, 1–333.

European Commission (EC), 2004. Corrigendum to Commission Directive 2004/73/EC of 29 April 2004 adapting to technical progress for the 29th time Council Directive 67/548/EEC on the approximation of the laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. Official Journal of the European Union L 152, 30 April 2004, 3–310.

Freeberg, F.E., Griffith, J.F., Bruce, R.D., Bay, P.H.S., 1984. Correlation of animal test methods with human experience for household products. J. Toxic. Cut. Ocul. Toxicol. 1 (3), 53–64.

Freeberg, F.E., Hooker, D.T., Griffith, J.F., 1986a. Correlation of animal test methods with human experience for household products: an update. J. Toxic. Cut. Ocul. Toxicol. 5 (2), 115–123.

Freeberg, F.E., Nixon, G.A., Reer, P.J., Weaver, J.E., Bruce, R.D., Griffith, J.F., Sanders III, L.W., 1986b. Human and rabbit eye responses to chemical insult. Fund. Appl. Toxic. 7, 626–634.

Ghassemi, A., Sauers, L.H., Bruner, L.H., Reer, P.J., Hall, R.H. (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.

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.

Heywood, R., James, L.W., 1978. Towards objectivity in the assessment of eye irritation. J. Soc. Cosmet. Chem. 29, 25.

Interagency Coordination Committee on the Validation of Alternative Methods (ICCVAM), 2006. *In vitro* test methods for detecting ocular corrosives and severe irritants. Background review document for isolated chicken eye (ICE) test. <www.iccvam.niehs.nih.gov/docs/ocutox_docs/ocubrd/ice/iceappA.pdf>.

Jacobs, G., Martens, M., De Beer, J., 1987. Selecting optimal dosage volumes for eye irritation tests in the rabbit. J. Toxicol. Cut. Ocul. Toxicol. 6, 109–116.

Jester, J.V., 2006. Extent of corneal injury as a biomarker for hazard assessment and the development of alternative models to the Draize rabbit eye test. Cutan. Ocul. Toxicol. 25 (1), 41–54.

Kay, J.H., Calandra, J.C., 1962. Interpretation of eye irritation tests. J. Soc. Cosmet. Chem. 13, 281.

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. Reg. Toxicol. Pharm. 36 (1), 106–117.

Mishima, S., Gasset, A., Klyce, S.D., Baum, J.L., 1966. Determination of tear volume and tear flow. Invest. Ophthalmol. 5, 264–276.

Organization for Economic Co-Operation and Development (OECD), 2002. Guideline for the testing of chemicals: acute eye irritation/corrosion. Guideline 405, adopted 24 April 2002.

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 (3), 291–296.

Prinsen, M.K., Koëter, H.B.W.M., 1993. Justification of the enucleated eye test with eyes from slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits. Food Chem. Toxicol. 31, 69–76.

Prinsen, M.K., 2006. The Draize eye test and *in vitro* alternatives: a left-handed marriage? Toxicol. In Vitro 20 (1), 78–81.

Prinsen M.K., 2004. Chicken enucleated eye test. INVITTOX protocol. <http://evcam-sis.jrc.it/invittox/published/indexed_80.html>.

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.

Swanston, D.W., 1985. Assessment of the validity of animal techniques in eye irritation testing. Food Chem. Toxic. 23 (2), 169–173.
United Nations Economic Commission for Europe (UNECE), 2003. Globally harmonized system of classification and labelling of chemicals (GHS). <http://www.unece.org/trans/danger/publi/ghs/ghs_rev00/00files_e.html>.
United States Environmental Protection Agency (USEPA), 1998. Health effects. Test guidelines OPPTS 870.2400. Acute eye irritation. EPA 712–C–98–195. <http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/EPA/EPA_870_2400.pdf> (accessed May 2008).
Weil, C.S., Scala, R.A., 1971. Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. Toxicol. Appl. Pharmacol. 19, 276–360.
Worth, A., Balls, M., 2002. Local toxicity: acute dermal and ocular effects. In: Worth, A., Balls, M. (Eds.), Alternative (non-animal) Methods for Chemicals Testing: Current Status and Future Prospects. A Report Prepared by ECVAM and the ECVAM Working Group on Chemicals, vol. 30. ATLA, p. 125 (Suppl. 1).
York, M., Steiling, W., 1996. A critical review of the assessment of eye irritation potential using the Draize rabbit eye test. J. Appl. Pharmac. 18, 233–240.

Histopathology as an additional observation in the Isolated Chicken Eye Test

Menk K. Prinsen, M.K. , M.E.I. Schipper, M.V.W. Wijnands.
Histopathology in the isolated chicken eye test and comparison of different stainings of the cornea.

Toxicology in Vitro 25 (2011) 1475-1479

6

Toxicology in Vitro 25 (2011) 1475–1479

Contents lists available at ScienceDirect

Toxicology in Vitro

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

Histopathology in the isolated chicken eye test and comparison of different stainings of the cornea

Menk K. Prinsen [a,*], M.E.I. Schipper [b], M.V.W. Wijnands [a]

[a] *TNO Triskelion, Toxicology and Applied Pharmacology Department, P.O. Box 844, 3700 AV Zeist, The Netherlands*
[b] *University Medical Centre Utrecht, Department of Pathology, Utrecht, The Netherlands*

ARTICLE INFO

Article history:
Received 14 December 2010
Accepted 29 April 2011
Available online 7 May 2011

Keywords:
ICE
Staining method
Histopathology
Isolated eye
PAS staining
Depth-of-Injury
Eye irritation *in vitro*
Alternatives

ABSTRACT

The isolated chicken eye (ICE) test, developed at our Institute, is accepted by the OECD for identification of severe eye irritants. The OECD ICE Guideline (No. 438) encourages preservation of the treated eyes for possible histopathology of the cornea, which is believed to strengthen evidence of absence or presence of irritation and to help clarify borderline effects by assessment of the corneal Depth-of-Injury. Histopathology of the cornea in addition to the normal slit-lamp microscope assessment of corneal effects has already been performed routinely in ICE tests at our Institute, using two standard stainings (H&E and PAS). In this study, three other stainings (AZAN, EVG and Trichrome), more specific for collagen-rich membranes such as basement- and Bowman's membranes were examined with corneas exposed to four model compounds ranging from non- to severely irritating (corrosive). PAS appeared to be the superior staining method. Surprisingly, the well-known eye corrosive sodium hydroxide (NaOH, solid) did not visibly compromise the integrity of Bowman's or the basement membrane. Based on our experience, histopathology of the treated cornea is confirmative in relation to the standard assessment of eye irritation by slit-lamp observation in the ICE and in certain cases can help to evaluate borderline effects. Besides establishing the depth of injury, additional investigation of corneal limbal stem cell damage after chemical exposure might be appropriate to determine reversibility or irreversibility of eye effects.

1. Introduction

The isolated chicken eye (ICE) test, developed at our Institute, is one of two validated alternative methods accepted by the OECD for identification of severe irritants (OECD, 2009; Prinsen and Koëter, 1993; Prinsen, 1996). The OECD ICE Test Guideline No. 438 encourages preservation of the treated eyes in order to perform histopathology. Histopathology of the cornea is believed to strengthen evidence (absence or presence) of eye irritation and to help clarify borderline effects, especially those that are expected to be reversible or irreversible. The corneal Depth-of-Injury approach was introduced as an additional tool to more precisely determine the extent of initial corneal injury (Maurer et al., 2002; Jester, 2006; Jester et al., 2010; Scott et al., 2010). This approach is based on histopathology performed by light microscopy and *in vivo* confocal microscopy of rabbit corneas and by using biomarkers of cell death and viability. In general, it is believed that with non- or (mild) irritants the effects are limited to the epithelium of the cornea, while moderate to severe irritants also affect the deeper layers of the cornea such as stroma and endothelium. In their well thought-out publication, Maurer et al. recommended that any *ex vivo* or *in vitro* replacement of the rabbit eye irritation test should meet the following criteria: (1) assessment of injury should be three-dimensional, as injury is a three-dimensional process, (2) extent of injury may be assessed by extent of cytotoxicity within the cornea, (3) assessment of injury to epithelium, stroma and endothelium, (4) differentiate injury that is diffuse from injury that is focally extensive, and (5) assessment of injury at different time points. The ICE test including histopathology meets by far these pre-requisites and, moreover, histopathology of the cornea in addition to the normal slit-lamp microscope assessment of corneal effects is already performed in ICE tests at our Institute for more than ten years. Most of these tests are performed for sponsors and the results are confidential. However, an article concerning the performance of the ICE test with household cleaning products and including histopathology has been published (Schutte et al., 2009).

From the perspective of reversibility/irreversibility, damage not only to the epithelium, stroma and endothelium of the cornea, but also to the other structures present in the cornea, such as basement membrane or Bowman's- and to a lesser extent Descemet's

Abbreviations: AZAN, Azocarmine & Aniline; EU-CLP, European Union-Classification, Labelling and Packaging; EVG, Elastic Van Gieson; H&E, haematoxylin & eosin; ICE, isolated chicken eye; OECD, Organisation for Economic Co-operation and Development; PAS, periodic acid-schiff; TNO, Toegepast Natuurwetenschappelijk Onderzoek (Organization for Applied Scientific Research); Trichrome, Masson's trichrome; UN-GHS, United Nations-Globally Harmonised System.

* Corresponding author.
E-mail address: menk.prinsen@tno.triskelion.nl (M.K. Prinsen).

0887-2333/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tiv.2011.04.028

membranes, are important to evaluate. After all, extracellular matrices, such as the basement membrane, are the essential framework for wound healing (Wagoner, 1997).

Initially, the traditional haematoxylin & eosin (H&E) staining was used for the histopathological evaluation of the cornea. In an attempt to further improve the evaluation of the cornea, the Periodic Acid-Schiff (PAS) staining was applied. This turned out to provide a colour spectrum with more contrast. The various specific structures of the cornea became better discernable, which was helpful during microscopic evaluation of the corneal lesions. Therefore, it was decided to use PAS as the standard staining method in the ICE test from then on. Three other stainings, known to be more specific for collagen-rich membranes, were examined and compared to the H&E and PAS stainings. Chicken corneas were obtained from standard ICE tests with substances classified according to the usual classification systems, i.e. UN-GHS (2007) and EU-CLP (2008); (formerly EC criteria for labelling of EC, 1993). Three substances represented the categories non-classified (NC), irritating (Cat2 sub-divided into 2A and 2B in the UN-GHS classification system) and severely irritating (Cat1), and one represented a borderline case between Cat2 and Cat1. The focus of this investigation was on the quality and applicability of the different staining techniques in relation to the morphology of the various cell structures and the pathology. Publications on the *ex vivo* histopathology of the cornea in the open literature are scarce or the work is confidential. Statements with respect to the corneal Depth-of-Injury theory and histopathological evaluation of the cornea in the ICE test are based on the many years of experience in this field at our Institute.

2. Materials and methods

2.1. Test substances

The four materials and their regulatory classifications selected were:

- physiological saline; non-classified/negative control (Eurovet, Bladel, The Netherlands)
- liquid surfactant containing cleaning product; Category 2/2A/R36 (source: confidential)
- liquid surfactant containing cleaning product; Category 2/2A/R36 borderline Category 1/R41 (source: confidential)
- NaOH, solid (purity > 97%); Category 1/R41 (Sigma–Aldrich, Germany).

Chicken corneas treated with these materials were obtained from routinely performed ICE tests, which constituted assessment of corneal swelling, opacity, fluorescein retention by damaged epithelial cells using the Haag–Streit slit-lamp microscope over a 4 h period and after a 10 s treatment with the test substance (OECD TG 438, 2009; Prinsen and Koëter, 1993; Prinsen, 1996). On the basis of the severity (maximum mean score of three eyes) of the observed findings for corneal swelling, corneal opacity and fluorescein retention, the effects were divided into four classes, viz. I = no effect; II = slight effect; III = moderate effect; IV = severe effect. The final irritation classification is determined by the combination of the three classes obtained for the three endpoints (corneal swelling, corneal opacity and fluorescein retention) into predefined classification schemes (Prinsen and Koëter, 1993; Prinsen, 1996). In addition, to allow for numerical ranking and comparison an Irritation Index was calculated. This index is based on the addition of the maximum mean scores obtained for the parameters according to the following formula: Irritation Index = maximum mean corneal swelling + maximum mean opacity (×20) + mean fluorescein score (×20). The factor of 20 is included to give equal weight to the scores obtained for opacity and fluorescein retention in the index compared to the maximum swelling possible (ca. 60%).

Table 2
PAS staining of the cornea.

Step	Treatment of the slides
1	Deparaffinization
2	Periodic acid 0.5% for 10 min
3	Rinsing in water for 5 min
4	Rinsing in aqua dest. two times for 1 min
5	Schiff reagent for 30 min
6	Rinsing in water for 30 min
7	Haematoxyline for 30 s
8	Short rinse in water
9	Dehydration, xylol, malinol
10	Pertex mounting medium

2.2. Preservation of the cornea

Our experience with histopathology of the chicken cornea showed that specific fixatives, e.g. Davidsons, often suggested for fixation of eyes appeared not to be necessary. The treated corneas (eyes) were collected in a neutral aqueous phosphate-buffered 4% solution of formaldehyde at termination of the ICE test, i.e. 4 h after treatment. For that purpose, the eyes were first cut in half with a scalpel just behind the level of the lens and through the vitreous body. The half with the cornea and lens was placed in a glass container with approximately 20 ml of formalin. After fixation for at least 24 h, the tissue was trimmed with scissors in such a way that a thin piece containing the entire cornea and the adjacent sclera were embedded in paraffin wax. Longitudinal serial slides (sec-

Table 1
Slit-lamp examination: maximum mean scores for corneal swelling, opacity and fluorescein retention, irritation categories assigned, Irritation Index, and regulatory classifications.

Test material	Maximum mean score for			Irritation class[1]	Irritation Index[2]	Classifications (UN-GHS[3]/EU-CLP[4]/EC-standards[5])
	Swelling %	Opacity	Fluorescein retention			
Saline	0	0.0	0.0	I; I; I	0	NC/NC/NC
Cleaning product 1	11	2.2	2.0	II; III; III	94	Category 2A/Category 2/R36
Cleaning product 2	18[6]	3.0	2.0	II; IV; III	118	Category 2A[7]/Category 2[7]/R36[7]
NaOH, solid	44	4.0	3.0	IV; IV; IV	184	Category 1/Category 1/R41

[1] I = no effect; II = slight effect; III = moderate effect; IV = severe effect.
[2] Irritation Index = maximum mean corneal swelling + maximum mean opacity (×20) + mean fluorescein score (×20).
[3] NC = not classified; Category 2B = mild irritant; Category 2A = irritant; Category 1 = irreversible effects on the eye/serious damage to the eye.
[4] NC = not classified; Category 2 = Irritating to eyes; Category 1 = irreversible effects on the eye/serious damage to the eye.
[5] NC = not classified; R36 = Irritating to eyes; R41 = risk of serious damage to eyes. EC-standards as published in the Official Journal of the European Communities, L 110 A, Volume 36, 4 May 1993.
[6] Wrinkling of the epithelium.
[7] Considered borderline with Category 1 and R41 because of the severe opacity and wrinkling of the epithelium.

tioned at 5 μM) were prepared from the central area of the cornea and further processed with the five different stainings.

2.3. Staining methods

In general, the directions given in the manual AFIP Laboratory Methods in Histotechnology (Prophet 1992) were followed for the H&E, the PAS, the Masson's trichrome (Trichrome), the Azocarmine & Aniline (AZAN) and the Elastic Van Gieson (EVG) stainings. Because PAS is used as the standard staining in the ICE test, its procedure is given in more detail (Table 2).

2.4. Histopathology

The stained slides were evaluated by the pathologist with respect to:

- general quality of the morphology of the epithelium, stroma and endothelium;
- visibility of the basement membrane;
- visibility of Bowman's membrane.

Evaluation of Descemet's membrane was not included because damage to this deepest situated membrane is not considered to result in borderline effects, but in severe, irreversible effects.

The routine semi-quantitative microscopic evaluation of PAS stained corneas was done by the pathologist according to the criteria shown in Table 3. The criteria for semi-quantitative microscopic evaluation were set after evaluating a large number of corneas ranging from unexposed to being exposed to severely irritating or corrosive materials.

3. Results

3.1. ICE test

Table 1 presents the summary results of the four materials assessed by slit-lamp observation in the ICE test.

3.2. Quality of the morphology

The results (Table 4 and Figures in Supplement) showed that general morphology could be assessed adequately with H&E and PAS and less so with Trichrome, AZAN and EVG. Bowman's membrane was visible with all stainings, whereas the basement membrane was visible with PAS only. H&E staining resulted in overall good quality, except for the visibility of the basement membrane, which was poor. PAS staining resulted in overall good quality with good discrimination of all relevant structures. The overall quality of Trichrome, AZAN and EVG was limited. The epithelium was difficult to evaluate after staining with Trichrome or AZAN, being either very dark or very red, respectively. The EVG staining resulted in hardly discernable stromal nuclei. Bowman's membrane was visible with all stainings, whereas the basement membrane was best visible with PAS.

3.3. Histopathological changes

Table 5 presents the routine semi-quantitative microscopic evaluation of PAS stained corneas treated with the non-classified/negative control, the Category 2/2A, the Category 2/2A borderline Category 1 and Category 1 materials. The evaluation comprised effects on the epithelium (erosion, necrosis and/or vacuolation), integrity of Bowman's- and basement membranes, the stroma (disorder of fibres and presence of pyknotic nuclei), and endothelium (necrosis). As expected the negative control cornea showed no abnormalities. With the irritating (Cat2/Cat2A/R36) test material, the corneal effects were limited to the epithelium, i.e. only very slight erosion and very slight vacuolation of the epithelium. No abnormalities were observed in stroma or endothelium. With the

Table 3
Semi-quantitative evaluation isolated chicken eyes.

Epithelium	
Erosion	
Very slight	Few single cells up to the entire single superficial layer.
Slight	Up to 3 layers.
Moderate	Up to 50% of the epithelial layer is gone.
Severe	Epithelial layer is gone up to the basement membrane.
Necrosis	
(Normal	<10 necrotic cells)
Very slight	10–20 necrotic cells.
Slight	20–40 necrotic cells
Moderate	Many necrotic cells but<50% of the epithelial layer.
Severe	50–100% of the epithelial layer is necrotic.
Vacuolation	
Very slight	Single to few scattered cells.
Slight	Groups of vacuolated cells or single string of cells with small vacuoles.
Moderate	Up to 50% of the epithelium consists of vacuolated cells.
Severe	50–100% of the epithelium consists of vacuolated cells.
Stroma	
Pyknotic nuclei	
(Normal	<5 pyknotic nuclei)
Few	5–10 pyknotic nuclei
Several	>10 pyknotic nuclei
Additional terms	
Undulating	Epithelial layer including the basement membrane is wrinkled.
General	
Unless otherwise indicated, lesions are always diffuse. However, in 'diffuse' lesions the central part of the cornea is usually more affected than the peripheral part. This may be due to the fact that the test substance, if a non-viscous liquid, which is applied on the centre of the cornea, dilutes when it flows to the peripheral parts of the cornea. In contrast, lesions can be classified focal or multifocal if they are actually confined to certain spots. This may be observed when the test substance is a powder	

Table 4
General quality of the morphology and the visibility of the Bowman's membrane and basement membrane.

Staining	Quality of the morphology	Visibility Bowman's membrane	Visibility basement membrane	Remarks
H&E	Good	Good	Poor	None
PAS	Good	Good	Good	None
Trichrome	Moderate	Moderate	None	Epithelium very dark
AZAN	Moderate	Good	Poor	Epithelium very red
EVG	Moderate	Good	Poor	Faded nuclei stroma

Table 5
Routine semi-quantitative microscopic evaluation of the cornea in the ICE (PAS staining; Fig. b1–b5).

Test Material	Epithelium			Bowman's and basement membrane	Stroma			Endothelium
	Erosion	Necrosis	Vacuolation	Integrity	Disorder of fibres	Pyknotic nuclei		Necrosis
						outer region (adjacent to epithelium)	inner region (adjacent to endothelium)	
Control (saline)	0	0	0	Not compromised	0	0	0	0
Category 2A	Very slight	0	Very slight	Not compromised	0	0	0	0
Category 2A/1	0[†]	0	Severe	Not compromised[‡]	0	0	0	0
Category 1 (NaOH)	Severe	0	0	Not compromised	0	Several	0	0[*]

0 = No abnormalities; ½ = very slight; 1 = slight; 2 = moderate; 3 = severe
[†] Epithelial layer (partly) detached from stroma.
[‡] Undulating appearance of epithelial layer, including Bowman's membrane (observed as wrinkling by slit-lamp observation).
[*] Occasionally in ICE tests with NaOH as the positive control, pyknotic nuclei in the inner region of the stroma and necrosis of the endothelium are also observed.

borderline severely irritating (Cat2/Cat2A/R36-Cat1) test material, the corneal effects were also limited to the epithelium but more pronounced, i.e. severe vacuolation of the epithelium. In addition, the epithelial layer was (partly) detached from the epithelium and had an undulating appearance, which was also noticeable by the routine slit-lamp observation (observed as wrinkling of the epithelium). With the severely irritating (Cat1) test material, the corneal effects were limited to the epithelium and the upper part of the stroma, consisting of severe erosion of the epithelium and several pyknotic nuclei in the stroma. No abnormalities were observed in the lower stroma or endothelium. Other ICE studies including the routine testing of this positive control (NaOH, solid), regularly show microscopic changes such as pyknotic nuclei in the lower stroma and necrosis of the endothelium (unpublished data).

4. Discussion and conclusions

This study compared five different stainings of the chicken cornea exposed in the ICE test to substances with different eye irritating properties. PAS, followed by H&E, was clearly the more optimal and preferred staining for evaluation of histopathological changes of the cornea in ICE tests. Its staining was superior with respect to the quality of the morphology of the corneal structures and the visibility of the different membranes present in the cornea. Trichome, AZAN and EVG were less successful with respect to the staining of the critical parts of the (chicken) cornea. Solid NaOH, one of the most corrosive materials, did not visibly compromise the integrity of Bowman's or basement membrane. In other routine ICE tests with severe irritants or corrosives, including NaOH solid, which could show damage to the deeper region of the cornea, damage to Bowman's or basement membrane was also not noticed (unpublished data). It is not clear if this lack of visible damage to Bowman's or basement membrane also means that the functionality of these membranes is uncompromised. Nevertheless, the examination of Bowman's membrane in the ICE test might be of importance because chicken corneas are similar to primate and human corneas in that they possess a Bowman's membrane, whereas the rabbit cornea lacks a Bowman's membrane (Fowler et al., 2004).

In general, information on histopathology of the human cornea in relation to chemical exposure is almost absent, but an interesting observation was the fact that vacuolation of the corneal epithelium in humans can occur after exposure to certain chemicals, whereas this lesion did not occur in rabbit corneas (Morton, 1986). In the ICE test, vacuolation of the corneal epithelium is part of the assessment and regularly observed (Schutte et al., 2009; Tables 3 and 5 in this article).

Other information, i.e. communications with an ophthalmologist specialized in the human cornea and publications on chemical injuries of the human eye and stem cell therapy (Wagoner, 1997; Rama et al., 2010), indicate that apart from degree of penetration, corneal limbal stem cells play an important role in regeneration of the cornea after chemical insult. Recovery from injury or successful grafting of corneal tissue is only possible if the corneal limbal stem cell deficiency is not too extensive after chemical insult. "If limbal stem cell loss is complete, severe superficial pannus invariably occurs (often in association with stromal vascularization) resulting in complete conjunctivalization of the new ocular surface" (Wagoner 1997). Because of their rather superficial location (early contact) and the role they play in recovery of corneal injury, i.e. by renewed epithelialization, one would expect that establishing corneal limbal stem cell survival after exposure could be of equal or more importance as the postulated Depth of Injury theory (Maurer et al., 2002; Jester, 2006; Jester et al., 2010; Scott et al., 2010). Preliminary investigations with respect to immunostaining of the chicken corneal limbal stem cells with p63 as a stem cell marker are ongoing at TNO.

If repair (reversibility) of corneal damage is depending on the resultant of the degree of penetration and limbal stem cell survival after exposure, routine histopathology of the cornea may not be sufficient to determine reversibility or irreversibility of eye effects. Remarkably, for labelling and classification of materials – this constitutes the purpose of the majority, if not all, of the acute eye irritation tests – there is no difference in labelling of severe irritants with reversible eye effects and severe irritants with irreversible eye effects. They are all labelled with Category 1 (irreversible effects on the eye/serious damage to the eye).

Based on our experience, we are of the opinion that, in general, histopathology of the treated cornea is confirmative in relation to the standard and robust assessment of eye irritation (damage) by slit-lamp microscope in the ICE test. Furthermore, it can clarify certain morphological phenomena observed by slit-lamp observation, such as wrinkling of the cornea, and assist in the evaluation of borderline effects.

Acknowledgements

The authors thank Dr. A. van der Lelij, ophthalmologist of the University Medical Centre of Utrecht and Dr C.F. Kuper of TNO Quality of Life for their contribution to the discussion.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.tiv.2011.04.028.

References

EC (1993). Commission Directive 93/21/EEC of 27 April 1993 adapting to the technical progress for the 18th time Council Directive 67/548/EEC on the approximation of laws, regulations and administrative provisions relating to the classification, labelling and packaging of dangerous substances. Official Journal of the European Communities, L 110A, 4 May.

EU-CLP (2008). Regulation (EC) No 1272/2008 of the European Parliament and of the Council of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006.

Fowler, W.C., Daniel, H.C., Roberts, Bruce.C., Zarovnaya, Elena.L., Proia, Alan.D., 2004. A new paradigm for corneal wound healing research: the white leghorn chicken (*Gallus gallus domesticus*). Current Eye Research 28 (4), 241–250.

Jester, J.V., 2006. Extent of corneal injury as a biomarker for hazard assessment and the development of alternative models to the draize rabbit eye test. Cutaneous and Ocular Toxicology 25, 41–54.

Jester, J.V., Ling, Joseph., Harbell, John., 2010. Measuring depth of injury (DOI) in an isolated rabbit eye irritation test (IRE) using biomarkers of cell death and viability. Toxicology In Vitro 24, 597–604.

Maurer, J.K., Parker, R.D., Jester, J.V., 2002. Extent of initial corneal injury as the mechanistic basis of ocular irritation: key findings and recommendations for the development of alternative assays. Regulatory Toxicology and Pharmacology 36, 106–117.

Morton, W.M., 1986. Toxicology of the eye, vol. 1. Charles c. Thomas Publishers, Springfield, Illinois, USA, p. 9.

OECD guideline no. 438: "Isolated chicken eye test method for identifying ocular corrosives and severe irritants", adopted on 7 September 2009, Paris, France.

Prinsen, M.K., Koëter, H.B.W.M., 1993. Justification of the enucleated eye test with eyes of slaughterhouse animals as an alternative to the draize eye irritation test with rabbits. Food and Chemical Toxicology 31 (1), 69–76.

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 and Chemical Toxicology 34 (3), 291–296.

Prophet, E.B., Mills, B., Arrington, J.B., Sobin, L.H., 1992. Armed forces institute of pathology, laboratory methods in histotechnology. American Registry of Pathology, Washington DC, United States.

Rama, P., Matuska, S., Paganoni, G., Spinelli, A., De Luca, M., Pellegrini, G., 2010. Limbal stem-cell therapy and long-term corneal regeneration. The New England Journal of Medicine 363 (2), 147–155.

Schutte, K., Prinsen, M.K., McNamee, P.M., Roggeband, R., 2009. The isolated chicken eye test as a suitable *in vitro* method for determining the eye irritation potential of household cleaning products. Regulatory Toxicology and Pharmacology 54, 272–281.

Laurie, Scott. et al., 2010. A proposed eye irritation testing strategy to reduce and replace in vivo studies using Bottom-Up and Top-Down approaches. Toxicology In Vitro 24, 1–9.

UN-GHS, 2007. United Nations-Economic Commission for Europe (UN/ECE). 2003. Globally Harmonised System of Classification and Labelling of Chemicals (GHS). UN, New York and Geneva, 2007.

Wagoner, M.D., 1997. Chemical Injuries of the eye: current concepts in pathophysiology and therapy. Survey of Ophthalmology 41 (4), 275–313.

The international validation process of the Isolated Chicken Eye Test

Menk K. Prinsen.
An Evaluation of the OECD Proposal for the Harmonised Classification of Eye Irritants and Corrosives.
Report of ECVAM Workshop 34, Eye Irritation Testing: The Way Forward, Appendix 1.

ATLA 27, 72-77, 1999

7

Appendix 1

An Evaluation of the OECD Proposal for the Harmonised Classification of Eye Irritants and Corrosives

Menk K. Prinsen

TNO Nutrition and Food Research Institute, Division of Toxicology, 3700 AJ Zeist, The Netherlands

Summary — Classifications of eye irritation/corrosion were assigned to 59 of the chemicals used in the European Commission/British Home Office (EC/HO) validation study by applying both European Union (EU) criteria and the harmonised criteria proposed by the OECD. It was found that the application of the two classification systems to the 59 chemicals resulted in comparable classifications: all of the chemicals classified as R36 or R41 according to EU criteria were classified as Category B or Category A, respectively, according to the proposed OECD criteria. Only two of the 59 chemicals, ethanol and ethyl-2-methylacetoacetate, were classified differently by the two systems; they were unclassified by the EU system, and were classified as Category B by the OECD system. It is concluded that: a) the proposed OECD classification system is broadly equivalent to the EU classification system; and b) future validation studies on alternatives to the Draize test would benefit from the application of the OECD classification system to the test chemicals.

Introduction

The failure of recent validation studies (1, 2) to find a suitable replacement to the Draize test is partly a result of the statistical method chosen for evaluating the performance of the *in vitro* tests. The relevance of these methods was assessed by correlating the *in vitro* test scores with the Modified Maximum Average Draize Test Score (MMAS), which is problematic because the MMAS shows considerable variability, particularly in the middle of the irritancy range. The use of the MMAS as the *in vivo* endpoint in an international validation study could also be regarded as inappropriate on the grounds that most regulatory systems (for example, the European Union [EU], US Environmental Protection Agency, US Food and Drug Administration, and Canadian workplace systems) do not classify chemicals on the basis of their MMAS values, but according to their effects in individual tissues of the eye (conjunctiva, cornea and iris), taking into account the recovery from or irreversibility of these effects. However, it would be difficult to accommodate all of the different classification systems when conducting a validation study. Fortunately, this should soon be unnecessary, since an OECD proposal for the global harmonisation of criteria for the classification of eye irritants (3, 4) is in its final phase of acceptance. The aim of this study was to examine the effect of applying the OECD criteria to chemicals which have already been classified according to EU criteria (5).

Materials and Methods

The animal data for 59 of the European Commission/British Home Office (EC/HO) chemicals were taken from the ECETOC reference chemicals data bank (6). These data were used to classify the 59 chemicals according to both EU criteria (5) and the proposed OECD criteria (3, 4), which are summarised in Tables I and II, respectively. Chemicals which

Table I: European Union classification system for eye irritation/corrosion

	R36 (irritating to eyes)		R41 (risk of serious damage to eyes)[a]	
Effect	**Three animals[b]**	**Six animals[c]**	**Three animals[b]**	**Six animals[c]**
Corneal opacity	≥ 2.0, but < 3.0	≥ 2.0, but < 3.0	≥ 3.0	≥ 3.0
Iris lesion	≥ 1.0, but < 2.0	≥ 1.0, but ≤ 1.5	≥ 2.0	> 1.5
Conjuctival redness	≥ 2.5	≥ 2.5		
Conjunctival chemosis	≥ 2.0	≥ 2.0		

[a]*A classification of R41 is also assigned if one of the eye effects has not reversed at the end of the observation period in at least one animal.*

[b]*The classification is assigned if the mean tissue effect (averaged over the 24-hour, 48-hour and 72-hour periods) is greater than or equal to the threshold value in at least two of the three animals. In this study, the same criteria were applied if four rabbits were used.*

[c]*The classification is assigned if the mean tissue effect (averaged over the three periods and over the six animals) is greater than or equal to the threshold value.*

7

Table II: Proposed OECD classification system for eye irritation/corrosion

Effect	**Category B[a]**	**Category A[b]**
Corneal opacity	≥ 1.0	≥ 3.0
Iris lesion	≥ 1.0	> 1.5
Conjunctival redness	≥ 2.0	
Conjunctival chemosis	≥ 2.0	

[a]*All effects have to be reversible within 21 days of treatment. The subcategory of B1 can be used for chemicals considered to be mildly irritating to the eyes, i.e. chemicals whose eye effects are reversible within 7 days of treatment.*

[b]*Category B is also applicable if an eye effect has not reversed, or is expected to reverse, within 21 days of treatment in at least one animal.*

Table III: Classifications of eye irritation/corrosion for 59 chemicals obtained by applying European Union (EU) and OECD criteria

No.	Test chemical	*In vivo* data		Classification[a]	
		Reference[b]	MMAS	EU	OECD
1.	Sodium hydroxide (10%)	82	108	R41	A
2.	Benzalkonium chloride (10%)	186	108	R41	A
3.	Trichloroacetic acid (30%)	36	106	R41	A
4.	Cetylpyridium bromide (10%)	193	90	R41	A
5.	Cetylpyridium bromide (6%)	191	86	R41	A
6.	Benzalkonium chloride (5%)	184	84	R41	A
7.	Captan 90 concentrate	170	83	R41	A
8.	Chlorhexidine	231	82	R41	A
9.	Cyclohexanol	77	80	R41	A
10.	Quinacrine	230	82	R41	A
11.	Promethazine hydrochloride	229	72	R41	A
12.	Parafluoroaniline	105	70	R41	A
13.	Triton X-100 (10%)	207	69	R36	B
14.	Acetone	157	66	R36	B
15.	Hexanol	74	65	R36	B
16.	1-Naphthalene acetic acid, sodium salt	168	64	R41	A
17.	Sodium oxalate	147	61	R41	A
18.	Isobutanol	70	60	R36	B
19.	Imidazole	124	59	R41	A
20.	Sodium lauryl sulphate (15%)	176	59	R36	B
21.	2-Ethyl-1-hexanol	68	51	R36	B
22.	4-Carboxybenzaldehyde	79	50	R36	B
23.	Methyl ethyl ketone	155	50	R36	B
24.	Pyridine	123	48	R41	A
25.	1-Naphthalene acetic acid	166	47	R41	A
26.	Benzalkonium chloride (1%)	180/182	34/56	R41	A
27.	2,2-Dimethylbutanoic acid	34	45	R41	A
28.	γ-Butyrolactone	228	43	R36	B
29.	Thiourea	—[c]	—[c]	—[c]	—[c]
30.	Octanol	66	41	R36	B
31.	Methyl acetate	28	40	R36	B
32.	L-Aspartic acid	33	37	R36	B
33.	Benzoyl-L-tartaric acid	32	37	R41	A
34.	Triton X-100 (5%)	203/205	32/34	NI/R36	NI/B
35.	Potassium cyanate	146	31	R36	B

Table III: continued

No.	Test chemical	*In vivo* data: Reference[b]	*In vivo* data: MMAS	Classification[a]: EU	Classification[a]: OECD
36.	Isopropanol	64	30	R36	B
37.	Sodium perborate	144	30	R41	A
38.	Dibenzyl phosphate	161	30	R36	B
39.	2,5-Dimethylhexanediol	63	28	R41	A
40.	Methyl cyanoacetate	27	28	R36	B
41.	Sodium hydroxide (1%)	80	26	R36	B
42.	Ethanol	62	24	**NI**	**B**
43.	2,6-Dichlorobenzoyl chloride	49	24	R36	B
44.	Ammonium nitrate	143	18	R36	B
45.	Ethyl-2-methylacetoacetate	26	18	**NI**	**B**
46.	Sodium lauryl sulphate (3%)	174	16	NI	NI
47.	Ethyl acetate	24	15	NI	NI
48.	Maneb	164	14	R36	B
49.	Fomesafen (acid form)	163	14	NI	NI
50.	Tetraaminopyrimidine sulphate	122	10	NI	NI
51.	Toluene	101	9	NI	NI
52.	Butyl acetate	20	8	NI	NI
53.	Trichloroacetic acid (3%)	30	7	NI	NI
54.	Methyl isobutyl ketone	149	5	NI	NI
55.	Tween 20	201	4	NI	NI
56.	Ethyl trimethyl acetate	18	4	NI	NI
57.	Methylcyclopentane	138	4	NI	NI
58.	Cetylpyridinium bromide (0.1%)	187	3	NI	NI
59.	Glycerol	56	2	NI	NI
60.	Polyethylene glycol 400	195	0	NI	NI

[a]*NI = non-irritant; R36/category B = irritating to eyes; R41/category A = risk of serious damage to eyes.*

[b]*Data from reference 6.*

[c]*Acutely toxic, and therefore discarded from study.*

MMAS = Modified Maximum Average Draize Test Score.

Bold type denotes different classifications obtained by applying the EU and OECD systems.

could not be classified as irritant or corrosive to the eye were classified as non-irritant (NI).

In addition, the 59 chemicals were divided into three groups: MMAS ≤ 25; 25 < MMAS ≤ 59; and MMAS > 59. These cut-off values were arbitrarily chosen as means of classifying chemicals into three groups. This three-fold categorisation of chemicals was compared with the EU classification of chemicals (NI/R36/R41).

Results

The EU and OECD classifications for the 59 chemicals, ordered in terms of decreasing MMAS, are given in Table III.

R41 classifications

Twenty two of the 59 chemicals were classified as R41, of which sodium hydroxide (10%; chemical 1) and benzalkonium chloride (10%; chemical 2) had the highest MMAS values (108), whereas 2,5-dimethylhexanediol (chemical 39) had the lowest MMAS (28). The R41 classification for the latter chemical is due to the persistence of eye effects in at least one rabbit.

R36 classifications

Twenty chemicals were classified as R36, of which the non-ionic surfactant Triton X-100 (10%; chemical 13) had the highest MMAS (69) and the pesticide Maneb (chemical 48) had the lowest MMAS (14).

NI classifications

Seventeen chemicals were classified as NI; Triton X-100 (5%; chemical 34) had the highest MMAS (32) and polyethylene glycol 400 (chemical 60) had the lowest MMAS (0).

Harmonised OECD classifications

With the exception of two chemicals, all of the chemicals classified as R41 were also classified as Category A, and all of the chemicals classified as R36 were also classified as Category B. The two exceptions, ethanol and ethyl-2-methylacetoacetate, were classified as NI on the basis of EU criteria, but as Category B on the basis of the harmonised OECD criteria. The subcategory B1 (mildly irritating) was not taken into consideration. The difference in the classification of ethanol under the two systems is interesting, given that ethanol is used as the positive control in the bovine corneal opacity/permeability (BCOP) assay.

Application of MMAS cut-offs

The comparison between the EU classifications and the classifications obtained by applying the MMAS cut-offs of 25 and 59 is summarised in Table IV. The results show that the cut-off values cannot be used to classify chemicals reliably.

Conclusions

It is concluded that: a) the EU and proposed OECD systems for the classification of eye irritants/corrosives are broadly equivalent; b) the MMAS cut-offs of 25 and 59 are not appropriate for classifying chemicals according to the two systems; and c) future validation studies on alternatives to the Draize test would benefit from the use of classifications based on the proposed OECD harmonised system.

Acknowledgement

The author gratefully acknowledges Ingrid Gerner (BgVV, Berlin, Germany) for her

Table IV: Classification of eye irritants on the basis of Modified Maximum Average Draize Test Score (MMAS)

	European Union classification			
	NI	**R36**	**R41**	**Total**
MMAS ≤ 25	16	3	0	19
25 < MMAS ≤ 59	1	13	8	22
MMAS > 59	0	4	14	18
Total	17	20	22	59

NI = non-irritant; R36 = irritating to eyes; R41 = risk of serious damage of eyes.

constructive comments regarding this work.

References

1. Balls, M., Botham, P.A., Bruner, L.H. & Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology in Vitro* **9**, 871–929.
2. Brantom, P.G. *et al.* (1997). A summary report of the COLIPA international validation study on alternatives to the Draize rabbit eye irritation test. *Toxicology in Vitro* **11**, 141–179.
3. OECD (1996). *Final Report of the OECD Workshop on Harmonization of Validation and Acceptance Criteria for Alternative Toxicological Test Methods,* 62pp. Paris, France: OECD.
4. OECD (1998). *Revised Proposal for the Harmonization of Hazard Classification Based on Eye Irritation/Corrosion. ENV/MC/CHEM/HCL(98)5,* 10pp. Paris, France: OECD.
5. EC (1993). Commission Directive *93/21/EEC* of 27 April 1993 adapting to technical progress for the 18th time Council Directive *67/548/EEC* on the approximation of laws, regulations and administrative provisions relating to the classification, packaging and labelling of dangerous substances. *Official Journal of the European Communities* **L110A**, 1–86.
6. ECETOC (1998). *ECETOC Technical Report No. 48, Eye Irritation: Reference Chemicals Data Bank*, 2nd edn, 236pp. Brussels, Belgium: ECETOC.

Regulatory acceptance of the Isolated Chicken Eye Test

M.K. Prinsen.
The Draize Eye Test and *in vitro* alternatives; a left-handed marriage?

Toxicology In Vitro 20 (2006) 78-81

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Toxicology in Vitro 20 (2006) 78–81
Toxicology in Vitro
www.elsevier.com/locate/toxinvit

Discussion

The Draize Eye Test and in vitro alternatives; a left-handed marriage?

M.K. Prinsen *

TNO Quality of Life, Toxicology and Applied Pharmacology, P.O. Box 360, 3700 AJ, Zeist, Utrechtseweg 48, 3704 HE, The Netherlands

Received 20 June 2005; accepted 20 June 2005
Available online 1 August 2005

Keywords: Draize Eye Test; Validation; In vitro assay

1. Variability of the Draize Test

No other animal test like the Draize Eye Irritation Test has been as controversial to replacement by in vitro methods, while initially it was believed to be one of the 'simplest' animal tests to be replaced. Since the early 1980s numerous alternatives have been developed, with some being submitted to validation, but without finding a single test or set of tests for replacing the animal test. Why is this? For many of the alternatives, it soon became clear that the chosen test system had not enough relevance with respect to eye irritation as was hoped for. For instance, a test system measuring decreased sperm mobility/motility provides some information on cytotoxicity in general, but not specifically on toxicity to corneal or conjunctival tissue. The fact that the toxicity measured by the test system has to be translated to specific (rabbit) ocular toxicity is the basis for most of the problems encountered. Furthermore, the variability of the Draize Eye Test, especially in the middle range of irritancy adds to this problem. The factors contributing to this variability are meanwhile well-known and recognized by the scientific world. The variability is mainly caused by the subjective scoring by different observers and interlaboratory variability.

* Tel.: +31 30 69 44 558; fax: +31 30 69 60 264.
E-mail address: prinsen@voeding.tno.nl

0887-2333/$ - see front matter
doi:10.1016/j.tiv.2005.06.030

2. Exposure conditions in the Draize Test

What is not highlighted in the discussions so far, however, is surprisingly enough the conduct and course of the test itself, although several investigators have discussed the unrealistic exposure conditions of the Draize Eye Test, i.e., instillation in the conjunctival cul-de-sac of the rabbits eye, compared to potential human exposure (Freeberg et al., 1986; Roggeband et al., 2000).

For most routine acute and repeat toxicity tests, standard exposure times and/or delivery of dosage (orally, intravenously, etc.) are well-defined. In the dermal irritation test, for example, the entire dosage is held by a patch onto the skin for an exact period of time. In the eye irritation test, however, neither of these well-defined conditions exists. For liquids, pastes and solids, it is impossible to estimate how much and for how long the test substance stays in contact with the eye. For aqueous, non-viscous formulations the standard instillation of 0.1 ml in the conjunctival cul-de-sac of the rabbit and the holding of the eye-lids for 1 s, results in a rapid removal of the material within seconds/minutes through blinking with the nictitating membrane (third eye-lid) and grooming by the rabbit.

This contrasts with the situation for sticky pastes for example, which cannot be removed that easily. The most dramatic variation in contact time and dosage occurs with solids. Even if applied as a 0.1 ml equivalent (the content of the cul-de-sac), the actual amount of a powder/solid that stays in contact with the eye is unpredictable. More importantly, the contact time may vary from a couple of minutes to 24 h, because rinsing the eye is

not allowed before the 24-h reading (only recently changed to 1 h for solids in the 2002 update of OECD guideline no. 405).

3. Testing of solids and variability

From ethical and scientific points of view, it is unbelievable that this situation still exists. Having carried out the test since 1981, it became increasingly difficult for me to adhere to this non-rinsing practice. Unintentionally, I discovered that the problem could be solved by manipulation of the eye-lids of the rabbit at the 1 h observation time point in such a way that any remnants of the test material present could be removed without rinsing. This process was always recorded in our reports, but never resulted in any comment on this deviation from the guideline. It is striking how few reports on eye irritation even mention the presence of remnants of powders/solids in the eye at the 1-h and/or 24-h observation time points, whereas it should be a common finding. The enclosure of solid materials up to 24 h in the conjunctival cul-de-sac, sometimes in combination with mechanical damage, can have a devastating effect on the eye. In the case of poorly water-soluble solids with distinct cytotoxic properties, the entrapped solid can rapidly cause a considerable and increasing degree of swelling of the conjunctivae, making it even more difficult for the animal to remove the material. If, at the 1-h observation, the lower eye-lid is not pulled away far enough by the observer, it can stay unnoticed that a bulk of test material lays deeply hidden in the conjunctival cul-de-sac.

Often, this forced continuous exposure for the next 24 h results in a complete closure of the eye-lids by the abundant production of colloidal discharge which often forms a sealing crust. Upon opening these sealed eye-lids, purulent discharge, and other inflammatory debris are released. The degree of swelling of the conjunctivae can be sufficiently severe such that removal of any remains of the test substance is hardly possible anymore. In the majority of these cases, the eye is permanently damaged or can only be saved by applying special care, such as regular daily cleaning and rinsing of the eye and eye-lids, often including cutting off the eye-lashes to prevent further sealing. In general, keeping the eye-lids open is essential for the recovery process, otherwise the enclosed inflammatory exudate will further damage the cornea. If no further extensive remedial treatment is given to the animal, the described exposure conditions can easily cause an initial opacity score of 1 or 2 to develop into a score of 3 or 4. Also, the eye can become vulnerable to microbiological infection (the so-called secondary inflammatory process), causing initial mild to moderate effects during the first days after exposure developing into more severe and prolonged effects during the 21 day observation period.

Without doubt such events will have occurred in other laboratories in the past, and probably will continue to occur, even with application of the present 1-h rinsing protocol for solids now in place. The events described here are of course not typical for all solids. Many of the solids are inert and form an unharmful bulk which can easily be removed by the animal or the observer, or they are well water-soluble and have already disappeared at the 1-h observation time point. However, the overall problem makes the Draize Eye Test highly variable, even *before* the actual scoring of effects takes place. Therefore, even if the scoring could be made more objective and less variable, the scores recorded will still represent a large variation. To my knowledge, this important source of variability has never been discussed, while its implication for any validation of alternative in vitro methods is very important.

4. Draize Test results and validation

Does this mean that we cannot use the data from the Draize Eye Test for validation purposes at all? It seems logical to assume that non-irritating or severely irritating hydrophilic liquids and non-irritating solids produce reliable reactions in the Draize Eye Test. Extremely variable results, however, will be obtained with sticky pastes and solids in the moderate to severe range of irritancy and with hydrophobic solutions. Such data will be unsuitable for the use as benchmarking data in the validation of in vitro methods. Apart from that, the kind of entrapment of solids in the rabbit eye has little relevance, because it is highly unlikely to occur in humans, accidentally or intentionally. For that reason, the low-volume eye test (LVET) has been developed by the Procter and Gamble company (Griffith et al., 1980). In this test, one-tenth of the original volume of 0.1 ml is administered directly onto the cornea of the rabbit and this is believed to mimic human exposure more realistically.

From a safety standpoint, it is understandable that the Draize Eye Test is still required by regulatory agencies, mostly because of the perceived higher sensitivity of the rabbit eye compared to the human eye. However, this perception has more to do with these exaggerated exposure conditions rather than with specific ocular tissue sensitivity. In that light, it is praiseworthy that, several years ago, European regulatory agencies took the initiative to accept in vitro screening of severe eye irritants by using isolated eyes or corneas, or the Hen's-egg chorioallantoic membrane (HET-CAM assay).

Recently, the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) and the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM) of the United States of America initiated a programme to officially adopt these alternatives

into the US-guidelines for the screening of severe eye irritants. In this programme, the methods and the data available are evaluated by a panel composed of national and international experts. Although there was a general awareness amongst the panel members concerning the variability of the Draize Eye Test, the general attitude was still to attempt to fit in the in vitro methods with the Draize Eye Test, rather than to address the validity of the latter test. All emphasis is again on the statistical evaluation of the in vivo and in vitro data. It is true that the nature and quality of the in vivo and in vitro data was examined in more detail, but mostly with the intention of modifying/optimizing the in vitro assays, rather than questioning the relevance of the in vivo data. To some extent the flaws of the rabbit test were acknowledged, but this did not lead to any real changes to the conduct of the test itself, whereas in vitro methods are still judged against this "Gold Standard" and repeatedly forced to have their methodologies adapted to the rather unrealistic conditions of the Draize Eye Test.

To give an example, ICCVAM evaluated the data of the EC/HO validation study (Balls et al., 1995) in which four candidates (Isolated Chicken Eye test, Isolated Rabbit Eye test, HET-CAM assay, and Bovine Corneal Opacity and Permeability test) selected by ICCVAM participated. Twenty-two out of the 59 substances examined in this study were severe irritants and, for the purpose of selecting in vitro methods for screening severe irritants, the data of these compounds (tested by four labs per method) were useful. For the Isolated Chicken Eye test (ICE), the ICCVAM panel concluded that it could identify severe irritants but with a high false-negative rate, especially for solids. Of the 22 compounds, the ICE identified 13 as severely irritating, 4 as irritating and 5 as non, or borderline, irritating. The latter five compounds were all solids. Remarkably, most of the in vitro methods participating in the EC/HO study also did not identify these compounds as severe irritants. Instead of questioning the in vivo exposure conditions of solids, ICCVAM considered this to be a deficiency of the in vitro method. Therefore, ICCVAM recommended that the test method needed to be optimized with respect to the exposure conditions for solids. Considering the fact that we have to deal with in vivo exposures to solids ranging from a couple of minutes up to 24 h, then standardization of the Draize Eye Test would be an appropriate recommendation.

5. How to proceed?

All suggestions for optimization/modification—how ever well intended they may seem to be—are driven by the thought that they are needed because there is not sufficient correlation with the in vivo "Gold Standard" test. The fact that these are totally uncontrolled and non-standardized conditions in the in vivo test, which cannot be modeled accurately by any of the in vitro tests, seems to be ignored or of no concern to regulatory bodies or to validation bodies like ICCVAM and ECVAM (European Centre for the Validation of Alternative Methods). Until the problems with the Draize Test discussed in this paper are solved and taken account of, all efforts to validate in vitro tests as complete replacements for the in vivo test will be doomed to fail.

Since the first international (pilot) validation (Commission of the European Communities, 1991) of alternatives for eye irritation was started in 1988, numerous validations using optimized/modified/standardized in vitro protocols have been carried out without any substantial success. We seem to be caught in a vicious circle and, by now, after almost 18 years of validation, I think it is time to conclude that further attempts will be futile, if we keep on using "old" in vivo data or new data generated by the current protocol for comparison. In fact, since the very first validation, most of the in vitro tests have been used in practice for decision making by many companies and have been accepted in Europe for screening of severe irritants. Having carried out the Draize Eye Test since 1981 and applied an in vitro/ex vivo screen prior to any in vivo testing since 1992, my recommendation would be a multi-way approach, as follows:

(a) Immediate implementation in the guidelines (legislation) of the most current in vitro methods in the testing strategy for screening of severe irritants, following current EC practice (CA, 2002). Many contract or company laboratories already have extensive experience with the existing in vitro alternatives. Moreover, severe irritancy is not based on the in vitro screen alone but often confirmed by other tests, such as skin irritation/corrosion (in vivo and in vitro) and acute dermal toxicity. Also indications of the possible (severe) irritating properties of a compound are often known by the manufacturer. Furthermore, most new substances will be tested in a battery of acute toxicity tests, covering skin and eye irritation, oral, dermal and inhalatory toxicity and skin sensitization, which require a tiered decision-making process by the investigator with respect to dose and test concentration selection. For that purpose, it is always useful to know as early as possible if one is dealing with an irritating (reactive) compound or not.
(b) Internationally, the Draize Eye Test should be re-evaluated taking a more realistic procedure like the LVET into consideration. The exposure conditions should be standardized for liquids and solids, i.e., a fixed exposure time, amount and mode of instillation (directly onto the cornea instead of in

the conjunctival cul-de-sac). In the EC guidelines there is the provision that substances causing eye irritation may also be examined for the effect of rinsing after a fixed exposure time, but in practice this possibility seems not to be followed. For exceptional circumstances, like ocular therapeutics or pesticides, the non-rinsing protocol could be maintained because in daily practice rinsing the eyes after (accidental) exposure to pesticides by the user may not always be possible.

(c) Together with the immediate implementation of in vitro methods and standardization of the Draize Eye Test, the possibilities for a more mechanistically-based development and optimization of in vitro methods should be an ongoing process. The parallel testing mentioned under point (a) would also offer the unique possibility to further validate the in vitro methods for the non-severe irritating category of compounds, and to test any new (mechanistically-based) modification both in vitro and in vivo.

References

Balls, M., Botham, P.A., Bruner, L.H., Spielmann, H., 1995. The EC/HO international validation study on alternatives to the draize eye irritation test. Toxicology in Vitro 9 (6), 871–929.

Commission of the European Communities, 1991. Collaborative study on the evaluation of alternative methods to the eye irritation test. EC Document XI/632/91, V/E/1/131/91.

CA, 2002. Competent Authority meeting of the European Communities. Agreement on the use of positive results obtained in non validated in vitro eye irritation tests.

Freeberg, F.E., Nixon, G.A., Reer, P.J., Weaver, J.E., Bruce, R.D., Griffith, J.F., Sanders, L.W., 1986. Human and rabbit eye responses to chemical insult. Fundamental and Applied Toxicology 7, 626–634.

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. Toxicology and Applied Pharmacology 55, 501–513.

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 and Chemical Toxicology 38 (8), 727–734.

General Discussion

To be Published in Regulatory Toxicology and Pharmacology

In this thesis the process of development, validation and acceptance of an alternative method to the Draize eye irritation test in rabbits is critically analysed and recommendations are presented to improve this process. What seemed to be one of the *in vivo* tests "easiest" to be replaced by an alternative test method, turned out to be a test that could only be replaced after a validation process of more than three decades. The Draize eye test was considered easiest to replace because it concerned a single dose test with topical exposure to a distinct target organ (the eye) and because many research groups were already exploring different alternative *in vitro* systems for eye irritation. The present thesis describes, mostly chronologically, the development and optimisation, (in house) validation and application of the Isolated Eye Test and, in a broader perspective, the international validation and acceptance of this alternative test by regulatory authorities and agencies.
A considerable part of the discussion also deals with the *in vivo* Draize rabbit eye test itself, because its performance and the use of its results are considered to be the main obstacle for replacing one of the most controversial experimental animal tests within safety testing today.
Since the introduction of the alternative test with isolated eyes, two different isolated eye models exist for the same test, each having two different abbreviations, namely the Rabbit Enucleated Eye Test or the Isolated Rabbit Eye test (REET or IRE) and the Chicken Enucleated Eye Test or the Isolated Chicken Eye test (CEET or ICE). For the sake of clarity, the abbreviations IRE and ICE will be used throughout this discussion.

The Isolated Rabbit Eye Test

Chapter 2 describes the implementation and validation of the IRE during the period 1982-1985. During that time, the superfusion apparatus and eye-clamps, as described by Burton (1981), were improved by TNO:
1) applying other more convenient dimensions for the individual chambers of the apparatus,
2) a three-side instead of a two-side water-jacket surrounding the chambers, and
3) eye-clamps consisting of a lower and upper ring instead of arms with pins (Figure 1).

Figure 1. Modified eye clamp (left) and original eye clamp (right).

With the improved equipment, thirty-four compounds already tested *in vivo* at TNO were retested in the IRE test. The use of additional rabbits serving as eye donors was avoided by reusing rabbits which had already been submitted to *in vivo* eye irritation testing (unused control eye) or to *in vivo* skin irritation testing (both eyes), and which had to be sacrificed at termination of the *in vivo* study. The IRE results showed an overall correlation of 82% with the *in vivo* results, using a general classification scheme for the grading of *in vivo* and *in vitro* eye irritation (i.e. not, slight, moderate or severe irritant). Four compounds, overpredicted (slight or moderate instead of non-irritant) by the IRE, were all moderate to severe irritants in the *in vivo* skin irritation test.
A common physico-chemical characteristic of these compounds was their hydrophobicity. A general observation in the *in vivo* rabbit eye test was that hydrophobic compounds stayed in contact with the cornea (eye) for a relatively short period of time, because they mixed poorly with the tear film on the cornea and because the nictitating membrane (third eye-lid) acted as a wiper, and rapidly removed the compound from the eye. This was in distinct contrast to the skin irritation test where the compound was kept in contact with the skin with the aid of a patch and fixative tapes for at least 4 hours. The presence of a third eye-lid is an example of a condition very specific to the *in vivo* rabbit eye test influencing the exposure. Hence this can cause a problem with respect to the validation of alternative methods since an alternative method is not able to mimic the presence of a third eye-lid (which is also an irrelevant condition with respect to humans). The issue of physico-chemical properties of the compounds and the interpretation of the *in vivo*/*in vitro* results in relation to the practical aspects of the *in vivo* rabbit eye test is further described in the discussion of **Chapters 7** and **8**. On the basis of the results with the 34 compounds, the IRE was considered to be a sensitive and useful test system for the identification of eye irritants. Negative *in vitro* results should be confirmed only in case of expected regular eye contact and in a maximum of three rabbits. At that time six rabbits was the usual number in the Draize eye test in order to comply with the US guidelines, although the first OECD guideline on eye irritation (OECD, 1981) recommended the use of 3 rabbits, which could mean a considerable reduction in the use of rabbits for eye irritation testing.

The use of slaughterhouse animals

An important milestone in the development of the isolated eye test is described in **Chapter 3.** Although rabbits are available as eye-donor for the isolated eye test in contract research organisations (CROs) executing routine eye- and skin irritation studies, the scientific community considered this dependency on laboratory animals a serious shortcoming. The isolated eye test would still be associated with animal experimentation and organisations not using laboratory rabbits could have difficulties in obtaining rabbit eyes. Therefore, the use of slaughterhouse animals, such as the cow, pig or chicken, as eye-donor was investigated. Firstly, the possibility to obtain eyes from abattoirs was explored and, secondly, possible practical limitations of the type of eye in relation to the experimental set-up were identified. On the basis of these explorations, the cow and pig were rejected as suitable eye-donor, although the latter was expected to be the most suitable candidate on the basis of its physiological resemblance with the human eye. The rejection of the cow eye was based on the following observations: 1) the irregular supply and variable origin of cows and, related to that, 2) the insufficient quality of the cornea in too many cases. In the case of the pig eye, the collection of suitable eyes by the investigator at the process line of the slaughterhouse appeared difficult, but doable. However, an important feature of the pig's eye (also applying to the cow's eye) was unfavourable for its use in the isolated eye test, namely its corneal thickness. The pig's cornea is relatively thick (600-700 µm; Faber, 2008) when compared to the cornea of the rabbit (ca 400 µm; Chan, 1983 and Li, 1997), but quite comparable to the thickness of the human cornea (ca 530 µm; Doughty, 2000 and Fowler, 2004). The cow's cornea easily doubles that thickness (900-1100 µm; OECD TG 437). Although the pig eye is better comparable to the human eye, the isolated eye test has to produce matching results with the Draize rabbit eye test in order to be accepted as an alternative. Hence, an eye (cornea) that matches closest to the rabbit eye and not to that of the human eye was needed. The chicken provided such an eye, i.e. its dimensions and corneal thickness (ca 400 µm; Fowler 2004) are similar to the rabbit cornea. Another point of concern was the baseline corneal thickness of the pig which was in the upper scale of the thickness measurement device and recording of increased thickness to its full extend after treatment with moderate to severe irritants appeared not possible. Other possibilities for measuring the corneal thickness in the higher range were not available for the Haag-Streit microscope in use. Ultrasonic pachymetry instead of the mechanical (non-invasive) measurement device was an option but the procedure required touching the cornea with a probe. In case of irritating substances, touching the cornea could result in additional damage and was therefore considered unsuitable.

The epithelium of the cornea is the first barrier against (chemical) insult and as such its thickness (number of layers of epithelial cells) is of importance. One of the most used alternative tests for eye irritation is the BCOP, which uses excised and isolated corneas of the cow. Because of its corneal thickness (not used as a parameter in the BCOP), the exposure time needed to elicit an irritation response is 10 minutes (for liquids) or 4 hours (for solids) in the BCOP. In the ICE and IRE test, only a 10-second contact

period is needed to elicit a relevant irritation response. A relevant response means a response comparable to that observed in the *in vivo* rabbit test. Apart from a comparable thickness to the rabbit cornea and, to a certain extent to the human cornea, the chicken cornea has another feature, which might be an advantage to predict (human) eye irritation, i.e. a well-developed Bowman's membrane (Fowler 2004). This membrane, which is in between the epithelium and stroma, is also well developed in the human cornea, but poorly developed in the rabbit cornea. In the discussion of **Chapter 6**, more information on Bowman's membrane in relation to histopathology performed in the ICE is presented.

With respect to the availability of eyes, attaining chicken eyes appeared relatively easy by simply collecting heads just after the sedation of the animals at the process line, transporting them to the laboratory and enucleating them from the heads within 2 hours after sacrifice. This time period is sufficient to guarantee corneas meeting the acceptance criteria for testing, i.e. no or very slight opacity, no or very slight fluorescein staining, and a corneal thickness in the normal range. The daily processing of thousands of chickens at the slaughterhouse guaranteed a constant supply of suitable eyes (corneas). Since 1981 we have visited the same slaughterhouse (v.d. Bor, Nijkerkerveen) for our supply and only during a period of about 6 months in 2003 could eyes not be obtained because of an Avian flu break-out (H7N7 variant) in a large part of the Netherlands. Even then, eyes could be obtained from another slaughterhouse outside the affected region (van Miert, Breukelen).

The use of chicken eyes was evaluated by testing the 21 reference chemicals previously tested in the first (pilot) EC validation study of the IRE (EC, 1991) which were selected to be representative of currently used industrial chemicals of different chemical classes and ranging from non-irritant to severe irritant. Describing the criteria and scoring system of the *in vitro* corneal effects together with a Prediction Model (PM) for matching the EC scheme for classification and labelling of compounds (see Table 3 of the introduction) was an important step forward. The development of the prediction scheme was primarily a theoretical exercise based on the range of physiological responses observed in the ICE (corneal swelling, opacity and fluorescein retention of damage epithelial cells). Because the corneal effects determined in the ICE have a direct relationship with the *in vivo* response (e.g. *in vitro* opacity for prediction of *in vivo* opacity), a PM could be established based on the magnitude/range of the effects and not by using an empirically-derived mathematical algorithm to translate an *in vitro* effect to an *in vivo* effect. For instance, in the HET-CAM assay lysis of blood vessels of the chorioallantois membrane is measured as time (seconds) of first occurrence (Luepke, 1985). The number of seconds recorded cannot be translated directly to an *in vivo* effect, which is considered a serious limitation of the method. A PM could be designed only after computer calculations of data obtained for a relevant number of compounds by using the *in vivo* MMAS (Modified Maximum Average Score) as the sole parameter for the *in vivo* test. The computer calculation was based on a mathematical formula or conversion algorithm resulting in a single *in vitro* irritation index score comparable to

the *in vivo* MMAS. Computer calculations were also needed with the BCOP assay, using a cornea, because the opacity is measured as a reduced light transmission value and the epithelial damage as the amount of fluorescein penetrated through the cornea, leading to an *In Vitro* Irritation Score (IVIS = mean opacity value + (15 x mean permeability OD490 value)).
The ICE followed a theoretical approach not using *in vitro* data to be compared to the *in vivo* MMAS. By knowing the ranges of the *in vitro* ICE responses (swelling, opacity and fluorescein retention), cut-off values were chosen to identify different categories of effect, viz. a non (Category I), slight (Category II), moderate (Category III) or severe (Category IV) effect. Thus after testing a compound, three categories were established, i.e. one for swelling, one for opacity and one for fluorescein retention.
The assignment of the final irritation classification to a non-irritant, (slight or moderate) irritant or severe irritant was obtained by the combination of these three categories. Again, a theoretical and weighted division of the different combinations possible was made for each final classification. For instance, at the low end of the ICE classification system the combination of I/I/I is a non-irritant and at the high end the combination of IV/IV/IV is a severe irritant. The combinations possible and respective irritation classifications are shown in Table 1.

Table 1. ICE *in vitro* classification system.

General classification	UN-GHS classification[a]	Combinations of categories
Not irritating	Not classified	3 x I 2 x I, 1 x II
Slightly irritating	2B : *Mild irritant/causes eye irritation*	2 x II, 1 x I 3 x II 2 x II, 1 x III 1 x I, 1 x II, 1 x III
Moderately irritating	2A : *Irritant/causes eye irritation*	2 x III, 1 x I 2 x III, 1 x II 3 x III 1 x IV, 2 x I 1 x IV, 2 x II 1 x IV, 2 x III 1 x IV, 1 x III, 1 x II
Severely irritating	1 : *Irreversible effects on the eye/ serious damage to the eye*	2 x IV, 1 x I 2 x IV, 1 x II 3 x IV

[a] Globally Harmonised System of Classification and Labelling of Chemicals (UN-GHS). UN, New York and Geneva, 2007

Often the PM of other alternatives had to be revised after more compounds had been tested by the in vitro test or additional PM's were especially designed for certain categories of compounds. With the ICE, the criteria system for scoring the effects was never changed, while the classification system has been modified twice, i.e. once to accommodate the introduction of the UN-GHS classification system and secondly after adoption as an OECD guideline for non-irritants (OECD 438, 2013).
The use of the in vivo MMAS in establishing the PM proved to be less ideal than thought. The EC and UN-GHS systems do not use the MMAS for determination of the irritation classification of a compound. Instead, the individual in vivo tissue scores are used. Evaluation of the use of the MMAS in the validation study of the European Commission and the British Home Office (EC/HO study) showed a poor correlation with these classification systems (Prinsen, 1999). This is further presented in the summary and discussion of **Chapter 7**.
On the basis of the ICE study with 21 reference materials (Prinsen and Koëter, 1993) the conclusions were that: 1) although the ICE does not assess conjunctival damage, its sensitivity to predict ocular damage is not reduced, 2) the ICE correctly predicted the EC classifications of the 21 chemicals and 3) the ICE fitted in the previously updated EC B.5 and OECD 405 guidelines regarding acute eye irritation/corrosion now including recommendations to use alternatives for the prescreening or positive identification of strong eye irritants.

The Isolated Chicken Eye Test in practice

The results described in the previous paragraph demonstrated that the ICE test was ready to act as a prescreen for the Draize eye test in rabbits. Apart from fundamental research, TNO also performed routine contract toxicity tests with a regular flow of acute toxicity testing (acute skin/eye irritation, acute oral/dermal toxicity and skin sensitization) which provided an excellent platform to incorporate the ICE as a prescreen for the *in vivo* Draize eye irritation test. Acute toxicity tests are predominantly carried out for regulatory purposes and conducting an *in vitro* test prior to the *in vivo* test would mean additional costs for the sponsor. As a consequence, the sponsor could refuse inclusion of the ICE, whereas "parallel" testing of compounds in the ICE and the *in vivo* Draize eye test would provide valuable information on the performance of the ICE. Moreover, it would introduce the ICE to sponsors and regulatory authorities, enhancing the chance for international acceptance, and for participation in validation programs that were starting to be initiated. Therefore, it was decided to include the ICE in addition to the routine Draize eye test without extra costs for the sponsor.
In **Chapter 4,** the results are presented of this "parallel" testing of 44 compounds. These compounds were considered to be a relevant cross-section of compounds (raw chemicals, finished products and formulations) routinely produced by the (chemical) industry, and to provide evidence for the use of the ICE in the assessment of eye irritation for regulatory purposes. Instead of only comparing single *in vivo* and *in vitro* irritation index scores, as was the common practice in validation, the individual components

used for calculation of the index score were also analysed. This was a recommendation of the United States Interagency Regulatory Alternatives Group (IRAG) made during a workshop on eye irritation testing in Washington DC, in 1993 (Scala, 1997). Fourteen different *in vivo* scores were derived from each of the 44 *in vivo* tests, covering time scores (1-hr, 24 hr, 1-72 hr, days to recovery), index scores (MAS and total eye score), and individual tissue cores (cornea, area of cornea involved, iris, cornea and iris combined, conjunctival redness, conjunctival swelling, conjunctival discharge, conjunctival scores combined). The *in vivo* critical scores were compared to the critical scores of the ICE test, namely the scores for corneal swelling, corneal opacity, fluorescein retention of damaged corneal epithelium and an index score (combination of the three parameters). The overall correlations found for the *in vivo* scores with the ICE *in vitro* scores were 0.90 (index score), 0.91 (corneal swelling), 0.86 (corneal opacity) and 0.82 (fluorescein retention). The correlation between the *in vivo* conjunctival scores and the ICE scores were 0.92 (index score), 0.92 (corneal swelling), 0.93 (corneal opacity) and 0.86 (fluorescein retention). These correlations substantiated the conclusion made by Burton (1972) that a relationship exists between the *in vivo* conjunctival damage and the corneal scores of the isolated eye test. Moreover, in ophthalmology the term ocular surface was introduced to emphasize the potential interdependence of the epithelium of the cornea and the epithelium of the conjunctivae (Thoft, 1977 and Wagoner, 1997). "Subsequent clinical and research insights of Thoft and others provided compelling evidence of the functional relationships between these two adjacent cell populations". Furthermore, the proven relationship and high correlation between the critical scores of the ICE test and the Draize eye test demonstrated that the test is a relevant alternative to eye irritation, and that applying regulatory irritation classification systems is "just" a matter of choosing the appropriate threshold limits belonging to the different irritation classes. This was supported by the *in vivo* and *in vitro* EC classifications obtained for the 44 compounds. Overall, it was concluded that the ICE provided a practical prescreen for the Draize rabbit eye test and that only mild to moderate irritants in the ICE, generally showing the highest sensitivity to inter- and intra-laboratory variability, should be confirmed in the rabbit eye test. Eighteen years later, the OECD adopted this conclusion. Furthermore, the "parallel" testing showed the ICE test to be robust in the sense that the practical aspects are not complicated and relatively easy to control, i.e. a saline drip is sufficient to maintain the eyes in good condition, and all compounds, regardless the physico-chemical properties, can be assayed.

The in-house repeatability of the ICE was assessed to be adequate during the Reference Standard Validation of *in vitro* tests sponsored by ECVAM (Brantom, 2000). Two reference compounds for the group of siloxanes (decamethylcyclopentasiloxane and cyclohexylamino-functional PDMS) and two for the group of surfactants (Triton X-500 5% and cetylpyridium bromide 6%), representing non-irritants, Category 2B and 2A, and Category 1 compounds, were tested five times each on different occasions.

The publication of the results for the first 44 compounds did not result in termination

of the "parallel" testing program. The main reason for continuing the "parallel" testing was that it was considered unethical to perform any toxicity test on live animals without prior information on the reactivity of the test compound using a "relevant" biological structure such as the cornea. In those days, it was common practice to start different acute toxicity tests with a new compound almost simultaneously and, if different study directors were involved, often without consulting each other about the specific results of their studies, whereas the result of an acute irritation test would have influenced or helped their decision concerning the study design to be followed. It became apparent that the ICE fitted well in a tiered approach for acute toxicity testing. The results of the ICE provided not only information on eye irritation, but also gave information that could help to optimally design the other acute toxicity tests. If the ICE test showed severe irritancy, the *in vivo* eye irritation test was waived and the skin irritation test was initiated with one rabbit only. Important decisions for the conduct of the acute oral and dermal toxicity tests in rats could also be made on the basis of the outcome of the ICE. In most cases, these studies were started as a limit study with the highest dose level of 2000 mg/kg body weight. In case of severely irritating or corrosive compounds, the local effects on the stomach or skin could lead to severe suffering or even mortality of the animals. When the ICE test showed severe effects, dosing of high levels or high concentrations of corrosive compounds could be avoided. When the ICE showed no or negligible signs of irritation (cytotoxicity), the decision to perform the acute oral and dermal tests with the highest dose level or a high test concentration could be better justified.

It was not until 2001 that the OECD adopted the use of results of any other *in vitro* toxicity test on a compound in order to determine start levels to be used for the *in vivo* acute oral toxicity test in rats (OECD 423, 2001).

For skin sensitization tests, this approach reduced the number of animals necessary for testing. At that time, the standard test for skin sensitization was the Guinea Pig Maximization test (GPMT), requiring up to 40 animals for the main test and 6-9 animals for the preliminary skin irritation test. The preliminary skin irritation test was needed to establish the appropriate (maximum tolerable) concentrations for the various phases of the study (i.e. the intradermal and topical induction and the topical challenge). Normally ranges covering concentrations from 1% up to 100% had to be investigated. When the ICE showed no or negligible irritancy, the range to be examined could be limited to only 100% and one lower concentration, which in practice meant that only 3 and not 6-9 animals were needed. At present, the GPMT is replaced by the Local Lymph Node Assay (LLNA, OECD TG 429, 2010), and this guideline also mentions the use of results of any other *in vitro* toxicity test on a substance as an aid in dose selection.

In retrospect it can be concluded that the testing strategy was successful for the majority of compounds submitted for testing, i) identification of severe eye irritants without further *in vivo* testing, ii) tiered testing of skin irritants/corrosives, iii) determination of acceptable (non-severe) dose levels of corrosive compounds in acute oral

and dermal toxicity testing, and iv) reduction of the number of animals used in the preliminary irritation experiment of sensitization studies.

"Parallel" *in vivo* and *in vitro* eye irritation testing was continued with another 50 compounds, meaning that each compound was first tested in the ICE and, in case of non-severe irritancy, directly followed by an *in vivo* rabbit test, both in full compliance with the OECD principles of Good Laboratory Practice (GLP). These results were submitted to organizations dedicated to the validation of alternative non-animal tests, such as ECVAM and ICCVAM. Because the performance of the ICE was considered sufficiently established after the "parallel" testing of 94 compounds, another approach was introduced to make optimal use of the ICE at low extra costs. First a non-GLP ICE test with only one eye was carried out and depending on the outcome either a full GLP ICE test (in case of severe irritancy) or an *in vivo* rabbit eye test (in case of non-severe irritancy) was carried out. This procedure was followed until the ICE OECD guideline 438 was adopted to include the identification of non-irritants (OECD, 2013). From then on, only compounds identified by the ICE as irritating (Category 2) need to be tested in the *in vivo* eye irritation test in rabbits.

The Isolated Chicken Eye Test as a stand-alone test

The ICE was increasingly being used as a stand-alone test by companies, which pursued a non-animal safety testing strategy of their products. One of the major international household and personal care companies, the Procter & Gamble Company (P&G), adopted the ICE for the eye irritation screening of their products. A variety of products and research formulations, ranging from hand or machine dish-wash products, household or industrial laundry products, household cleaning products, cosmetics and hair dyes, were submitted for testing in the ICE (**Chapter 5**). This was a boost for the application of the ICE and further development of the method. In the past, P&G had heavily invested in the introduction of a modification of the Draize eye test, because the exposure conditions of the standard Draize eye test were considered unrealistic and exaggerated compared to (recorded) accidental human exposure. Under the standard conditions of the Draize eye test a considerable number of their products, especially detergents in powder form, had to be labelled as severe eye irritants. Their modification presented as an alternative method, the Low Volume Eye Test (LVET; Griffith,1980), uses only one tenth of the dose volume of the test material administered directly onto the cornea instead of instilling it in the lower conjunctival sac (Figure 2). P&G employed the LVET since 1980 and the results correlated well with their database on human responses

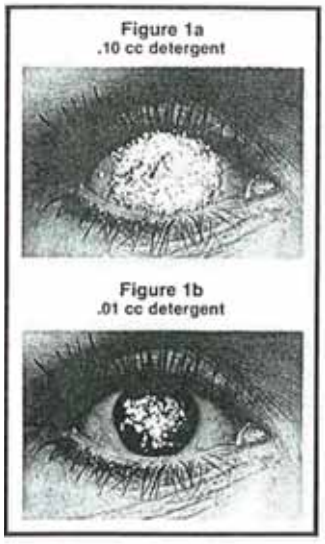


Figure 2. Dose volume of 0.1 mL versus 0.01 mL (Griffith, 1987).

to accidental exposure and to results of clinical studies (Griffith, 1980 and 1987). Although less burdening for the animals, the LVET was still an animal test and in view of their non-animal testing policy they decided to explore the suitability of the ICE for their products.

Twenty products (8 liquids and 12 powders) representing acidic or alkaline cleaners, dishwashing liquids or powder, powder detergents and raw materials of these products were tested in the ICE. In addition to the ICE standard application of 30 µl or 30 mg, which is relatively equivalent to the amount used in the Draize eye test, volumes of 10 and 3 µl and amounts of 10 and 3 mg were used to examine if this provided better correlation with the LVET results. Furthermore, histopathology of the treated corneas sampled at the end of the observation period was now included. The inclusion of histopathology was one of the recommendations made by Maurer (2002), based on their microscopic examination of the cornea in LVET studies. They suggested a direct correlation between the depth of injury in the cornea and possible recovery of eye lesions. Since measurement of recovery is not possible in the ICE (and in all other alternatives), it could be an important improvement of the method and helpful in the argumentation for absence or presence of microscopic lesions in the ICE that could predict (ir)reversibility. The outcome of this study (Schutte, 2009) was that, in general, the results of the ICE using the standard volume of 30 µl or mg were in line with or more conservative than the LVET in terms of classification. The level of overprediction found in the ICE was expected and considered acceptable since the ICE was developed for prediction of the Draize eye test and not the LVET.

For P&G the ICE test was considered useful for several purposes, such as 1) EU/GHS classification and labelling of powder and liquid household cleaning products, 2) screening of candidate formulations, and 3) weight-of-evidence approach by determining the profile of new cleaning product formulations against benchmark products. A definite conclusion on the usefulness of histopathology in the ICE could not be made, but the data showed that assessment of the histopathological lesions in the various parts of the cornea was possible, enabling the application of the "Depth-of-Injury" theory of Maurer. The question remained if this theory established on the basis of an inflammatory process in the rabbit's eye (consisting of initial ocular injury and subsequent repair over days/weeks) correlated with the irritation process or damage in the ICE. The outcome of the *in vivo* Draize rabbit eye test is heavily influenced by several unpredictable factors or events which will not occur in the ICE test. These factors and their consequences for the final classification of compounds are part of the discussion of **Chapter 8**.

Optimizing histopathology in the Isolated Chicken Eye Test

For the histopathological observations in the ICE, as presented in **Chapter 6**, a quite basic and routine procedure of processing the eye/cornea is used. The eye is preserved in formalin at the end of the study, i.e. 4 hours after the 10-second exposure, and further processed into a paraffin block from which slides of the longitudinal section of the corneal center are prepared. This area was considered appropriate since the application

of the test compound was made at the center of the cornea and in general the slit-lamp observation showed confluent, homogenous corneal effects. In case of non-homogenous effects, such as focal spots with more severe opacity not present in the central area, these parts were also examined. The choice of the staining appeared an important issue. In first instance, the most common staining by Haematoxylin & Eosin (H&E) was used, which was later replaced by the Periodic Acid-Schiff (PAS) staining, which provided a much better discrimination of the different layers of the cornea. The three major layers of the cornea, i.e. epithelium, stroma and endothelium were well visible with both stainings. The other structures, such as basement membrane and Bowman's membrane (between epithelium and stroma) and Descemet's membrane (between stroma and endothelium), were not that visible with H&E, whereas PAS provided much better results. The integrity of the membranes was considered to play a role in the injury and recovery process of the cornea, and the visibility of these membranes by microscope was considered important for an adequate histopathological assessment by the pathologist. Therefore, other staining methods specifically targeting collagen-rich membranes were tested on corneas treated with compounds representing a non-irritant, two irritants and a severe irritant (corrosive). The microscopic examination focused on the basement and Bowman's membrane and not on Descemet's membrane (see Introduction, Figure 2), because damage to this membrane adjacent to the endothelium was considered to result in severe, irreversible effects and not borderline effects. Of the five stainings selected, i.e. H&E, PAS, Trichrome, AZAN (Azocarmine & aniline) and EVG (Elastic Van Gieson), PAS was clearly superior with respect to visibility of the membranes and the quality of the morphology of the various corneal structures. Moreover, the histopathological examinations provided interesting facts and insights. After severe corrosive damage to the cornea by sodium hydroxide, observed macroscopically through slit-lamp observation as maximum swelling and very severe opacity of the cornea, the basement and the Bowman's membranes appeared undamaged while effects were seen in the underlying stroma. Does this mean that the functionality of the membranes was not compromised or is light microscopy not able to detect such damage of the membranes? This observation led to the conclusion that depth of injury is not the only factor determining the seriousness of corneal injury. Personal communications with an ophthalmologist of the University Medical Centre of Utrecht, specialized in the cornea, and publications (Wagoner, 1997; Terry, 2002; Rama, 2010) indicated that in the clinic, emphasis is put on corneal opacity and corneal stem cell survival after (chemical) injury (see Table 2). With severe stem cell damage, recovery of the corneal damage by re-epithelialization of migrating stem cells from the corneal limbus is not possible. In that case, the recovery process will result in complete, irreversible conjunctivalization of the ocular surface. The stem cells are rather superficially located in the limbal region of the eye and as such involved in early contact with a topically applied compound. Therefore, the possibility of screening the viability of stem cells after chemical injury could be of value. However, explorations to stain stem cells of the chicken cornea by p63 immunostaining appeared unsuccessful.

Table 2. Hughes-Roper-Hall classification of chemical burns of the human eye (Kim, 2002).

Grade	Findings	Prognosis
I	Corneal epithelial damage; no limbal ischemia	Good
II	Cornea hazy but iris details seen; ischemia less than ⅓ of limbus	Good
III	Total loss of corneal epithelium; stromal haze blurring iris details; ischemia at ⅓ to ½ of limbus	Guarded
IV	Cornea opaque, obscuring view of iris or pupil; ischemia at more than ½ of limbus	Poor

Although the assessment of reversibility or irreversibility in the ICE is undoubtedly of value, especially for household and personal care companies, the EU/GHS classification and labelling of severe eye irritants does not make a distinction between these two categories; they are all classified as Category 1: "irreversible effects on the eye/serious damage to the eye". The use of additional histopathology in the ICE over the past ten years has proven that it is mainly confirmative of the results already obtained by the slit-lamp observation and, in some instances, can be used to support decision making in borderline cases between irritating and severely irritating compounds. In general, the histopathology results provided no reason for altering the irritation classification. In literature on the cornea, an interesting observation was that vacuolation of epithelial cells can occur in the human cornea after chemical insult, but it does not occur in the rabbit cornea (Grant, 1986). In the ICE, especially with detergent products, vacuolation of epithelial cells was regularly observed by the pathologist. ICE studies commissioned by the International Association for Soaps, Detergents and Maintenance Products (AISE) showed that vacuolation of epithelial cells, and in particular its location in the epithelial layer (top, mid or bottom region), may be reason for an upgrade of the classification to Category 1 in case of borderline severe corneal effects (Cazalle, 2014).

International Validation of the ICE (EC/HO study)

With the introduction of alternative methods for the Draize Eye Test in the mid-eighties, a certain wild-growth of alternative methods occurred and internationally (mostly) in western Europe clusters of specific methods could be seen. Roughly, the IRE/ICE was developed and practiced mainly by the UK and the Netherlands, the Bovine and Corneal Opacity and Permeability test (BCOP) by France, the Hen's Egg Test Chorioallantoic Membrane (HET-CAM) by Germany and the Neutral Red Uptake test (NRU) by several European countries. With the increasing interest in these methods and the need for regulatory acceptance it became apparent that a formal validation of the alternative methods was needed.

A first (pilot) validation of several alternative methods, including HET-CAM, IRE and NRU was commissioned by the EU in 1987 (the Collaborative study on the evaluation of alternative methods to the eye irritation test, 1991), which was the basis of one of the largest validation programs held in the early nineties, known as the EC/HO study.

In this comprehensive study, organized by the European Commission (EC) and the British Home Office (HO), nine methods including the ICE (see Introduction, Table 6) were each carried out by four laboratories testing 60 chemicals which represented different chemical classes and irritation potential (Balls, 1995). The performance of the ICE in the EC/HO study, for which TNO was the lead laboratory, and the validation process in general is part of the discussion of **Chapters 7 and 8**.
The chemicals were selected from the ECETOC database and were considered to have reliable *in vivo* eye irritation data. Basically that meant the tests having been performed under GLP conditions and in compliance with OECD TG 405 (1981). No other assessment with respect to the quality of the individual data was made. The outcome of the study was quite disappointing; none of the methods were capable of identifying the eye irritation potential of the compounds (maximum overall correlation ranged from 0.34 to 0.55). Breaking up the compounds into different categories such as liquids, solids, surfactants, non-irritants, severe irritants etc. did not improve the correlation significantly, although the group of surfactants showed the best results across all methods.
One of the reasons for this disappointing result was considered to be the use of the MMAS as the sole parameter derived from the *in vivo* data. This score, ranging from 0-110, is an average of the maximum individual tissue scores of the animals recorded for a compound. By using certain cut-off values for the MMAS, a compound is classified as a non-irritant (score 0-25), irritant (score 25-59) or severe irritant (score >59). Assessment of the entire process of ocular inflammation by a single index instead of using the *in vivo* data to its full potential appeared rather inadequate. For instance, the classification system used in the EC does not use a single irritation index score, but is based on individual tissue scores (i.e. for cornea, iris and conjunctivae separately) and/or the (ir)reversibility of these effects within 21 days (Table 3 of the Introduction). In 1998, the OECD published a proposal for the harmonization of hazard classification based on eye irritation/corrosion, which is comparable to the EC classification system, because it also uses the individual tissue scores separately (Table 4 of the introduction). However, slightly lower thresholds were used for classification as an irritant (Category 2) or severe irritant (Category 1), and additionally recovery of eye effects within 7 days were used to discriminate between a mild irritant (2B) and an irritant (2A). Because the MMAS is not used for regulatory classification, the impact of the EC and proposed OECD classification criteria on the compounds tested in the EC/HO study was investigated (**Chapter 7**). First of all, it was established that applying the two classification systems to the EC/HO compounds resulted in classifications that were comparable between the two systems, i.e. R36 compounds were also Category 2A/2B compounds, and R41 compounds were also Category 1 compounds. Subsequently, it was demonstrated that the MMAS cut-offs of 25 and 59 belonging to, respectively, irritant and severe irritant, were not appropriate for classification according to these two regulatory systems. Eight compounds with an MMAS lower than 59 were in fact severe irritants (R41/Category 1) and 4 compounds with a MMAS higher than 59 and 3 compounds

with a MMAS lower than 25 were irritants (R36/Category 2). One of the reasons was that an MMAS could be lower than 59 during the study while ocular effects persisted until day 21 which according to the classification system is reason for R41/Category 1 classification. The recommendation was that validation of alternatives would benefit from the use of classifications based on the proposed OECD harmonized system (later on adopted as the EU-CLP and UN/GHS classification systems). After this publication on the role of the MMAS in 1999, it was not until 2004 that a new initiative was undertaken to (re-)validate four of the alternative methods which were considered the most promising in the EC/HO study, especially for the screening of severe irritants.

International Validation of the ICE (ICCVAM study)

The validation was an initiative of ICCVAM and NICEATM (National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods) of the USA in collaboration with ECVAM of the EC and the methods selected were the ICE, the HET–CAM, the IRE test and the BCOP. An independent Expert Panel was established for each alternative method to determine the validation status of these methods. After a public request for data on the four methods in 2004, a public meeting was held in January 2005 at the National Institutes of Health (NIH), Bethesda, USA to assess the current validation status of the *in vitro* test methods proposed for identifying compounds that may cause serious eye damage and to develop recommendations for further validation. During that meeting, the ICE method was presented and questions of the panel members were answered. The nature of the questions and the attitude towards the validation procedure in general were disappointing. The focus was on the belief that the general poor correlation with the Draize Eye Test was due to shortcomings in the practical performance of the *in vitro* methods which should be improved, rather than critically addressing the validity and shortcomings of the Draize eye test and its consequences for the validation of *in vitro* alternatives. Emphasis was again put on the statistical evaluation of the *in vivo* and *in vitro* data. Following that meeting, a discussion paper was written by M.K. Prinsen, entitled: "The Draize Eye Test and *in vitro* alternatives; a left-handed marriage?" (**Chapter 8** of the thesis).

Another public request for *in vitro* data was made by ICCVAM in February 2005.
The data of 94 compounds "parallel" tested in the ICE and the Draize eye test (all data per individual animal or per chicken cornea) were submitted by M.K. Prinsen of TNO. In total ICCVAM obtained data of 174 compounds, i.e. previously tested by i) Prinsen and Koëter (1993; 21 compounds), ii) Prinsen (1996 and 2005, dataset of 94 compounds in total), and iii) Balls (1995; 59 compounds), which were compiled in a background review document (ICCVAM BRD, 2006). This time the MMAS was no longer used for reanalysis of the accuracy and reliability of the ICE, but regulatory classification criteria (UN-GHS, EC and EPA) were applied as was recommended previously (Prinsen, 1999). The panel evaluated the additional data submitted on the methods in 2005. ICCVAM and its Ocular Toxicity Working Group summarized the Expert Panel evaluation, the revised analyses, the public comments, and the comments of the Scientific Advisory

Committee on Alternative Toxicological Methods in a final Test Method Evaluation Report (ICCVAM, 2006).
The conclusion was that *"there are sufficient data to substantiate the use of the BCOP or the ICE test methods, with certain limitations, as screening tests to identify compounds as ocular corrosives and severe irritants in a tiered-testing strategy, using a weight-of-evidence approach, for regulatory hazard classification purposes".* The limitations of the ICE were the testing of alcohols based upon the false positive rate and of solids and surfactants based upon the false negative rates. The limitations for the BCOP were the testing of alcohols and ketones based upon the false positive rates and of solids based upon the false negative rate.
Although the acceptance as a screen for ocular corrosives and severe irritants was a success for the ICE, the way the data were used by ICCVAM for reanalyses and for adjustments in the experimental procedures of the ICE was considered debatable.
It was acknowledged that the *in vivo* rabbit eye test was subject to variability, but the *in vivo* data, however, were taken as absolutely accurate (the "Golden Standard") in predicting the eye irritation potential of a compound.
The *in vivo* irritation classifications were assigned if the study performance met with the criteria set by ICCVAM. One of the criteria (assessment of full reversibility of any eye effect) was the reason that the database (ICCVAM BRD of ICE: Appendix D, March 2006) contained many gaps compared to the original data submitted, whereas sufficient *in vivo* data were available for classification. For example, two compounds (2,2-dimethyl butanoic acid and p-fluoroaniline) identified in the EC/HO validation study as severe irritants (R41) on the basis of the individual *in vivo* data (ECETOC, Technical document no. 48: 2, June 1998) were rejected with the remark "study criteria not met". The two compounds were correctly identified as R41 by the ICE and by most other *in vitro* methods participating in the validation study. The *in vivo* classification was based on sound scientific judgment, but ICCVAM decided to exclude the compounds, because a 21-day observation period was not completed. The OECD/EC guidelines (at the time of testing) specified that the observation period should be long enough to evaluate the reversibility or irreversibility of the lesions. The six rabbits treated with 2,2-dimethyl butanoic acid still showed slight to severe corneal opacity and neovascularization of the cornea at 14 days after treatment. It was considered evident that these lesions would not have cleared within a 21-day observation period. Thus, the 14-day observation period applied was in agreement with the guidelines and should not have been a reason for discarding the results by ICCVAM. The same applied for p-fluoroaniline, which caused moderate to severe corneal opacity and iritis score 2 (highest score possible; no reaction to light, haemorrhage, gross destruction). The test was terminated on day 3, which is also in agreement with the guidelines which state that animals may be humanely sacrificed if the severity of the effects is considered too high.
Similar cases occurred in the "parallel" data set of the 94 compounds provided by Prinsen (2005). These cases also mainly concerned *in vivo* eye irritation studies that were terminated earlier than 21 days after treatment because of the severity of eye effects

or compounds lacking an *in vivo* eye irritation study because of proven skin corrosivity in the *in vivo* rabbit skin irritation test (10 compounds), that was performed immediately after the ICE had shown severe eye irritancy. In full agreement with the guidelines, the *in vivo* eye irritation test in rabbits was waived in these ten cases. The individual *in vivo* skin irritation data and the ICE data were provided to ICCVAM, but the 10 compounds remained excluded from the analyses because of "classification assigned on the basis of skin corrosion assay" (SC). Remarkably, ICCVAM claimed to apply the criteria for classification according to the EC (1993) and UN-GHS (2007), whereas both guidelines unambiguously state "corrosive to skin" as one of the criteria for classifying a compound as severely eye irritating.

Another remarkable conclusion in the ICCVAM document was the underperformance of surfactants by the ICE. This conclusion was primarily based on the 6 surfactants examined in the EC/HO validation study. The correlation percentages for surfactants in the EC/HO study tested by the four participants that performed the ICE were 72, 82, 83 and 76%, compared to an overall mean correlation percentage of 54%. In general, the chemical class of surfactants was best predicted by each of the nine *in vitro* methods participating in the EC/HO study. The fact the ICE has been employed by P&G for more than ten years for their product development (Schutte, 2009), the majority of which contains surfactants covering the whole spectrum of eye irritancy, is in contrast with ICCVAM's conclusion. More recently, member companies of the International Association for Soaps, Detergents and Maintenance Products (AISE) increasingly use the ICE for their (surfactant-containing) products (Cazelle, 2014).

With respect to the ICE study performance the ICCVAM expert panel identified two major issues: 1) the variability in swelling percentages obtained by the four laboratories performing the ICE, and 2) the use of only one control eye per experiment.

Variability in swelling percentages (high CVs)

The variation in swelling percentage was caused by the use of different pachymeters with different slit width settings. This was already intensively discussed by the Management Team of the EC/HO study, but was considered of no concern because the *in vivo* MMAS was compared to the critical scores of the ICE (i.e. the max. mean swelling%, max. mean opacity score and mean fluorescein score) and not to the regulatory irritation classifications. ICCVAM overlooked this fact and decided to apply the TNO ICE system for categorizing effects to the other three participating labs as well. This was a valid approach for the opacity and the fluorescein scores because the scoring is exactly the same for the four labs, but it could not be used integrally for the swelling %. The misconception of the variability in corneal swelling by the ICCVAM expert panel led to incorrect conclusions and recommendations. For example, centering lights needed to be installed on the optical pachymeter to improve the determination of corneal thickness by ensuring consistent central corneal thickness measurements across laboratories. The purpose of these lights in human ophthalmology is to guide the patient's eye to a fixed point and thus perform the reading at the centre of the cornea.

This is used because the subjects often (involuntarily) move their eyes making the (central) corneal thickness assessment difficult and variable. The chicken eye is isolated and fixed, so there is no movement at all. Therefore, the corneal thickness can be measured in a very accurate and reproducible way at the centre of the cornea without any additional aid.

Use of control eyes

ICCVAM decided to increase the number of negative control eyes from the usual one per experiment to three, because three was the accepted minimum of replicates in *in vitro* testing in general. The use of only one negative control eye has been employed and approved by all users of the isolated eye test (both with rabbit eyes and with chicken eyes) since the introduction of the method by Burton in 1981 and during the EC/HO study. The use of only one negative control eye is justified since the effects of the cornea treated with the test material are not assessed or evaluated in any way against the effects of the control eye. This is possible because, prior to testing, the quality and suitability of each cornea can be accurately assessed and, moreover, each cornea provides its own pre-dose baseline thickness/opacity/fluorescein control values. Furthermore, all compounds (liquids, pastes and solids) are tested neat and, therefore, effects of solvents need not to be examined. The purpose of the negative control eye is only to demonstrate the appropriateness of the general conditions in the chambers of the superfusion apparatus, i.e. the saline drip onto the cornea and chamber temperature, necessary to maintain corneas in the proper condition during the 6-hour test period. All ICE experiments used for the ICCVAM BRD were performed with one negative control, representing 354 independent test runs or replicates. These negative controls never showed any unusual effects during the 6-hour test period and adequately demonstrated the appropriateness of using only one negative control for the purpose of monitoring the general conditions of the test system. With the practical limitations to the maximum number of chambers that can be constructed in the superfusion apparatus, the two additional control chambers could be better used to examine reference compounds or vehicle controls. Very rarely, an experiment was cancelled because the eyes (corneas) collected proved to be of insufficient quality at the first screening immediately after dissection from the head. Not one experiment failed because the control eye showed corneal effects during the experiment. Other alternatives like the BCOP need to use three or more control corneas because pre-dose corneal observations, to assess their suitability for testing, are not possible. The BCOP control values at each observation time point are needed for subtraction from the values of the test corneas.
Following a public request for comments on the draft ICCVAM Test Method Evaluation Report: "Evaluation of the Current Validation Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants" (Federal Register notice 13513, March 2005), a document was submitted by M.K. Prinsen in which the above issues were extensively highlighted and discussed. It was emphasized that TNO had been involved in the international validation process of alternative test systems for eye

irritation since the very first introduction of these methods in the early eighties. Throughout that validation process, one of the lessons learned was the extreme importance of exchanging information between experts, especially concerning the practical aspects of the test methods. In that light, it was regrettable that ICCVAM declined the offer of TNO to organize and host a meeting for key experts of the ICCVAM panel to demonstrate and discuss the ICE test method prior to the expert meeting in Washington. The presence and interaction with the panel experts would have been invaluable for understanding the practical aspects of the test method and probably would have avoided considerable discussion about certain aspects of the method.
Upon a second public request for comments (Federal Register notice 43149, July 2005), the "*In Vitro* Ocular Toxicity Draft Background Review Document (BRD)" on the ICE test method was commented by M.K. Prinsen. This time the discussion concentrated on the exclusion of the ICE data of skin corrosive compounds and other data gaps in the BRD. ICCVAM was requested to revise the analyses with respect to the screening of severe irritants by inclusion of the ten cases of skin corrosive compounds. Concern was also expressed on the decision to pool the data of the various ICE (validation) studies for analyses without analyzing the individual studies separately. *In vivo* and *in vitro* data obtained in "parallel" are of a higher quality than ICE data compared to *in vivo* data obtained from literature, because in the latter case it is obtained by different observers, with different batches of the compound, under different laboratory conditions and often with only summarizing data reporting. Furthermore, reservations were made about the handling of the *in vivo* data, i.e. that no lessons were learned from previous validation studies, which made clear that another approach for validation was needed. The disadvantages and shortcomings of the *in vivo* rabbit eye test and their implications on the test results should be addressed first before starting the validation process. To initiate a dialogue, the discussion paper (**Chapter 8**) was added to the comments sent to ICCVAM.

The Draize Eye Test and *in vitro* alternatives; a left-handed marriage?

In **Chapter 8,** the difficulties and circumstances that are encountered when performing the Draize eye test in rabbits and how they influence the final irritation classification of a compound are summarized and discussed. Recognizing, understanding, and the correct appreciation of the *in vivo* test conditions and their effect on the results are crucial for the evaluation of the *in vitro* results obtained by the alternative test method. The *in vitro* test conditions can usually be fully controlled in contrast to the conditions of testing in the *in vivo* rabbit test. The *in vivo* results are used as the "Golden Standard" for all comparisons with the alternative tests. However, the shortcomings of this test should not be ignored, especially if one is aware of the nature and extend of these shortcomings. Up to now, in compliance of the rabbit test with the current guidelines was the criterion used in validation studies for the validity and acceptance of the *in vivo* data. Once that was established, the *in vivo* results are treated as the absolute

truth without any room for interpretation or expert judgment. This means that a compound is classified as a non-irritant, an irritant or a severe irritant based on the *in vivo* test without any kind of nuance or deliberation. There is general agreement that the *in vivo* test in rabbits is far from perfect, but the implications of the inconsistencies of the *in vivo* test for the validation of the alternatives were never taken into account. In fact, over the years an increasingly rigid attitude towards questioning the value of the *in vivo* eye irritation data can be noted. This can be considered the root of the problems encountered since the very first validation took place in 1989.
There is a common believe and acknowledgement that the *in vivo* rabbit eye test produces variable results due to the fact that different labs and observers are involved and data had been obtained over a very long period of time. However, this is only a small part of the problem. There are more serious reasons to consider the *in vivo* data in a critical way.

There are several important issues that play a crucial role in the outcome of the *in vivo* test:

1. The kind of exposure
By instillation of the compound in the conjunctival cul-de-sac (Figure 3), closing the eye-lids for one second and then releasing the rabbit, the exposure is undefined. It can be anything from minutes (liquids) to 24 hours (solids). No washing out of remnants from the conjunctival sac was allowed before 24 hours after treatment. Only after modification of the OECD guideline in 2002 was a wash-out allowed after one hour. Especially with poorly soluble/dissolving powders the results can be devastating if the powder is present for one hour, let alone for up to 24 hours. It should be noted that the ICCVAM validation was mainly with *in vivo* data obtained before 2002. Remarkably, Draize only mentions the testing of liquids, solutions and ointments and not the

Figure 3. Conjunctival cul-de-sac (left) and a dose-volume equivalent of 0.01 mL on the cornea (right) used in the LVET test, i.e. 10 times lower than used in the Draize eye test (TNO).

testing of solids. In fact, in his 14-page publication (Draize, 1944), only one sentence is dedicated to the actual test procedure for eye irritation testing. Overall, eye irritation was dealt with in a rather limited way when compared to his discussion on dermal toxicology and skin sensitization. One wonders what would have happened if Draize had extended his eye irritation investigations to the testing of solids.
These undefined exposure conditions are in contrast to the basic principles of toxicity testing. Moreover, this kind of exposure condition by placing such a large amount of compound in a retracted eye-lid will hardly occur in humans (see also Figure 2).
A well-defined and standardized exposure in toxicity testing is one of the pillars of hazard and risk assessment. For instance, in the acute skin irritation in rabbits, a semi-occlusive exposure of 0.5 ml to 6.25 cm2 of skin is applied for 4 hours under a patch and fixative tape. These are standardized conditions and remarkably the skin irritation test in rabbits has been fully replaced by alternatives since 2010 (OECD 431 and OECD 439).

2. Behaviour of the animals

The behaviour of rabbits after treatment may also differ considerably. After treatment the animal is immediately released and is placed in its home cage where it can move freely. Usually, they start grooming and/or scratching. One rabbit out of a group of 3 treated may do this excessively (Table 3; rabbit no. 1227061), while on the other end of the behavioural spectrum another animal may freeze and not react at all. Again these variations in behaviour add considerably to the variability of the results.

Table 3. Results of a three-animal *in vivo* test (TNO Triskelion report, 2012).

Animal No.	Corneal effects:		Iris	Conjunctivae		Discharge
	score	area		redness	swelling	
			1 HOUR			
1227061	0	4	1	1	1	3
1228067	0	4	0	1	1	1
1228045	0	4	0	1	1	1
			24 HOURS			
1227061	1	4	1	1	1	1
1228067	0	4	0	0	0	0
1228045	0	4	0	0	0	0
			48 HOURS			
1227061	1	4	1	1	1	1
1228067	0	4	0	0	0	0
1228045	0	4	0	0	0	0
			72 HOURS			
1227061	0	4	0	1	1	0
1228067	0	4	0	0	0	0
1228045	0	4	0	0	0	0
			6 DAYS			
1227061	0	4	0	0	0	0
1228067	0	4	0	0	0	0
1228045	0	4	0	0	0	0

3. Treatment of the eye post-exposure

When significant irritation occurs in an early stage, the treatment of the eye post-exposure highly determines the outcome for classification. The observation times after treatment are essentially the only moments that the animals are handled outside the cage. In case of a moderate eye irritant those time points are normally 1 hour, 24 hours, 48 hours, 72 hours, 7 days, 14 days and 21 days. In between, the animals are not handled except for a cage-side observation once a day. The enclosure of solid materials up to 24 h in the conjunctival cul-de-sac, sometimes in combination with mechanical damage, can have a devastating effect on the eye. In the case of poorly water-soluble solids with distinct cytotoxic properties, the entrapped solid can rapidly cause a considerable and increasing swelling of the conjunctivae, making it very difficult for the animal to remove the material. If, at the 1-hour observation, the lower eye-lid is not pulled away far enough by the observer, a bulk of test material deeply hidden in the conjunctival cul-de-sac may remain unnoticed. In most cases this continuous exposure for the next 24 hours results in a complete closure of the eye-lids by the abundant production of colloidal discharge which often forms a sealing crust (Figure 4). Upon opening the sealed eyelids, purulent discharge, and other inflammatory debris are released. If the animal (treated eye) is not receiving special care of the eye an otherwise irritating compound can easily become a severe one.

Figure 4. Production of colloidal discharge sealing the eye-lids (Wikepedia and TNO).

The swelling of the conjunctivae can be such that removal of the remains of the test compound is hardly possible. In the majority of these cases, the eye is permanently damaged or can only be saved by applying special care, such as regular daily cleaning and rinsing of the eye and eye-lids, often including cutting off the eye-lashes to prevent further sealing. This special care is not common practise in the Draize eye test nor is it mentioned in the guidelines, whereas it can certainly relieve the discomfort and pain experienced by the rabbit considerably. In general, keeping the eye-lids open is essential for the recovery process, otherwise the enclosed inflammatory exudate will further damage the cornea. If no further extensive remedial treatment is given to the animal, the exposure conditions described can lead to an opacity score of 3 or 4 instead of the initial score of 1 or 2. In these cases, recovery from these injuries has little or no

relevance for man. As with the exposure conditions, these kind of circumstances are not representative for the human situation. After accidental exposure, one will seek "immediate" care in case of ocular damage, and the victim will usually receive medical treatment, if required. The unrealistic exposure conditions in the Draize eye test impelled P&G to develop the *in vivo* rabbit Low Volume Eye Test (LVET) for their products. For instance, the testing of a dish wash detergent tab would result in dramatic ocular effects in the standard *in vivo* Draize eye test, because the tab is ground to a fine powder and instilled as a bulk in the conjunctival cul-de-sac of the rabbit and remains there for at least one hour (before 2002 up to 24 hours). Nobody would consider this as real exposure circumstances, nor will it occur in real life. In the ICE, the exaggerated test conditions can be mimicked by leaving the powder on the cornea for up to 60 seconds instead of the standard 10 seconds, but what relevance does it have? The LVET was designed to mimic the possible human exposure and was extensively used for household care products. Now P&G uses the ICE for their purposes because the test also mimics the possible human exposure more closely. In general, it was astonishing that both ECVAM and ICCVAM urged that the ICE method needed to be modified in order to mimic the extreme exposure conditions of the *in vivo* Draize eye test, rather than modifying the exposure (to solids) in the *in vivo* Draize eye test.

Another phenomenon that occasionally occurred in the Draize eye test is the development of a secondary infection following the eye effects caused by the compound (initial infection). In the past, the hygiene standards in the laboratories were not as high as currently, and the treated eye could be infected by the scratching/grooming of the animal with its paws. In addition to the inflammation caused by a compound, the eye is more vulnerable to microbiological infection, causing initial mild to moderate effects during the first days after exposure developing into more severe and prolonged effects during the 21 day observation period. An interesting example of such an event can be found in one of the compounds tested in the EC/HO study and which was also used in the ICCVAM validation of the ICE.
1-Napthaleneacetic acid was tested in six rabbits of which one rabbit showed very unusual persisting and increasing effects after day 7, compared to the eye effects observed in the other five rabbits (ECETOC, 1998). The effects of the five rabbits followed a pattern which is normally expected for the initially slight to moderate eye effects, i.e. gradually decreasing in severity after day 3 and followed by a complete recovery on day 14 or day 21. In the sixth rabbit, a similar pattern was observed until day 7, but thereafter the slight opacity observed increased to a moderate opacity on day 10 and finally a very severe opacity on days 14 and 21. This difference in the pattern of the eye effects is remarkable and most probably caused by a secondary infection in the eye of the animal. Based on the result in this rabbit, the compound was classified as Category 1, whereas the initial effects (24-72 hours) in the 6 rabbits would lead to a Category 2(A) classification. The ICE test also classified the compound as Category 2. It is remarkable that the OECD guideline 405 of 1981 already stated that "Care should

be taken in the interpretation of data to exclude irritation resulting from secondary infection". However, this issue was not addressed in the EC/HO validation or in the ICCVAM validation.

4. Observation/grading of eye effects

In the early days of validation of alternatives for eye irritation it was recognized that the variability could be high in the Draize eye test, and this was considered to be caused by subjective scoring by different observers and by interlaboratory variability (Weil and Scala, 1971; Lordo, 1999; Ohno, 1999). Unfortunately, the publications by Lordo and Ohno did not include the individual *in vivo* rabbit eye data which might have provided more insight in the underlying causes of the variation. Subjective scoring is indeed part of the problem but a large part of the variation presently ascribed to subjective scoring might in fact be caused by differences in animal behaviour, differences in exposure times, and absence (or presence) of post-treatment care.
For instance in the study of Weil and Scala, ethanol 95% caused a combined score (all tissues combined) of 2, 9, 22, 15, 38 and 110 in the 6 rabbits at the 72-hour observation time point in one of the labatories. A score of 110 is the maximum score possible. Amongst the 24 laboratories the median score for ethanol 95% ranged from 0 to 42. This cannot only be explained by subjective scoring.
The subjective nature of the observation definitely plays a role with compounds causing effects near the thresholds for classification (not classified/irritant and irritant/severe irritant). First of all there is the grading/scoring of the effects itself. The Atlas of eye effects of the FDA (1964) already gave rise to debate. For example the redness of the conjunctivae of the eye no. 6 of Plate 1 (Figure 5) is stated to be score 2 (moderate redness: more diffuse, deeper crimson red, individual vessels not easily discernible; see also Table 2 of the Introduction). Based on the grading of eye effects of all compounds tested at TNO since 1981, it should be the maximum score of 3 (severe redness: diffuse beefy red), because a more intense redness cannot be observed. Eye no. 3 of Plate 1 is presented as a case of redness score 1 (slight redness: vessels definitely injected above normal), whereas this would be a good example of score 2 for redness. Eye no. 2 of Plate 1 is more representative of a redness score 1 than of a normal eye (redness: vessels normal). The other plates of the Atlas contain more examples of grading that are considered questionable and subject to debate.

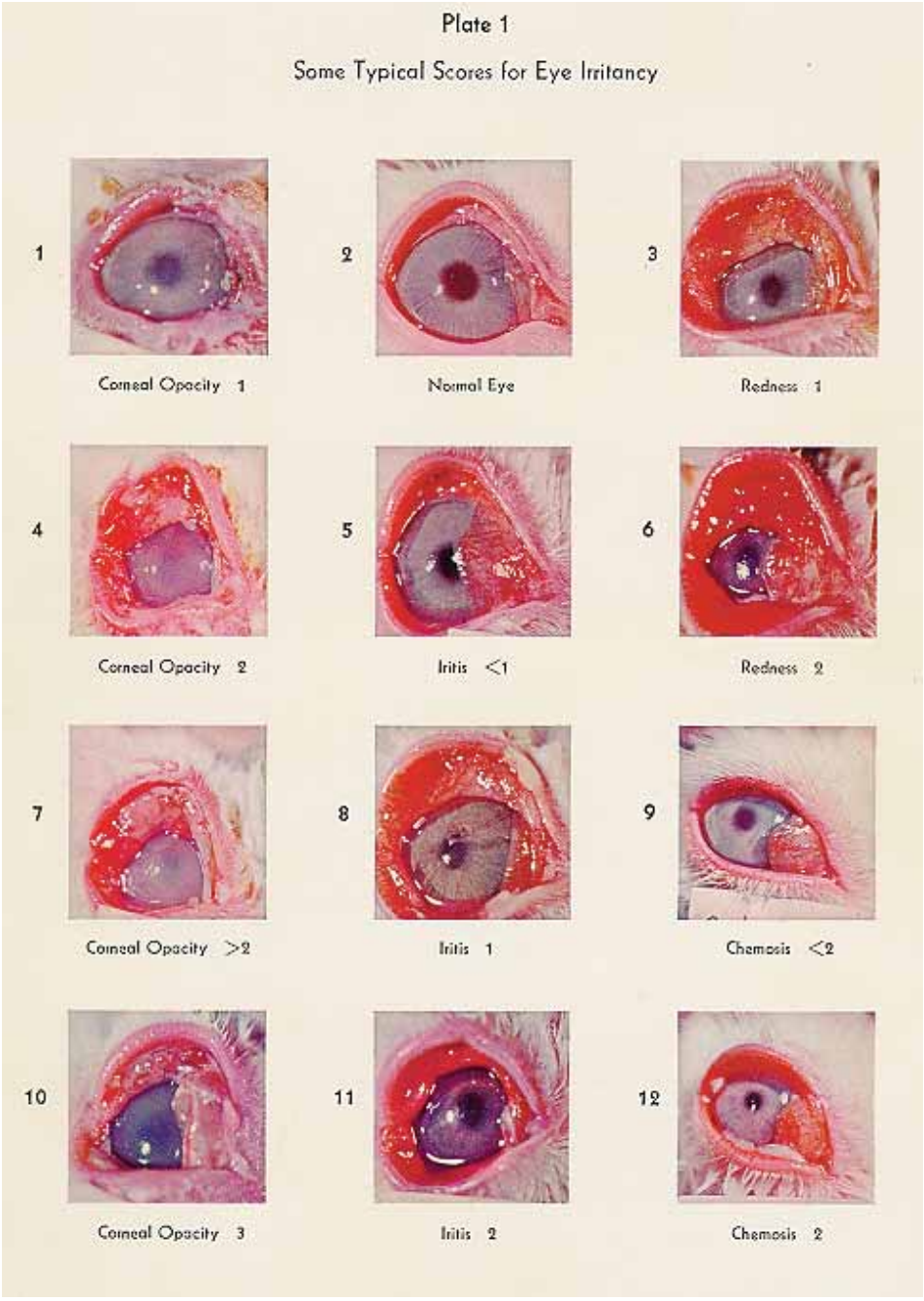


Figure 5. FDA guidance on scoring of ocular lesions; Plate 1 (FDA,1964).

Another issue concerning the subjective scoring is the decision the observer has to make in certain cases where the score of one animal at one time point can make the difference between, for instance, not classified and irritant (Category 2). For EU-CLP and UN-GHS classification, the individual tissue scores of each animal is first averaged over the 24-72 hour time points and next the average score of the two rabbits showing the highest score for a specific tissue determines classification or not. The threshold score for redness or for swelling of the conjunctivae is an overall average of 2.0 for classifying as a Category 2 compound. Table 4 shows a theoretical case where one of the 6 scores can make the difference between classifying or not. The score of 1.0 in Table 4 is assigned to animal 2 at the 72 hour time point, but can theoretically be at any of the 6 different places.

Table 4. Examples of conjunctival scores (redness or swelling); Category 2 (first table) versus Not Classified (second table).

Time point 24 Hours	
Animal 1	2.0
Animal 2	2.0
Time point 48 Hours	
Animal 1	2.0
Animal 2	2.0
Time point 72 Hours	
Animal 1	2.0
Animal 2	2.0

Time point 24 Hours	
Animal 1	2.0
Animal 2	2.0
Time point 48 Hours	
Animal 1	2.0
Animal 2	2.0
Time point 72 Hours	
Animal 1	2.0
Animal 2	1.0

That this is not merely a theoretical example is demonstrated by the scores obtained for Fomesafen, acid form (ECETOC, 1998) which was also part of the EC/HO and ICCVAM database (Table 5). The scores obtained for opacity are lacking just one score 1 for classification as Category 2. Also the scores for redness of the conjunctivae are very close to the threshold score of 2. The ICE test identified the compound as a Category 2B (UN-GHS criteria) or Category 2, but borderline to not classified (EU-CLP criteria). In the ICCVAM validation this result was considered to be a false positive.

Table 5. Ocular effects caused by Fomesafen, acid form (ECETOC, 1998).

***Fomesafen, Acid Form**

			Observation period (days)											
Animal No.	4		1 h	4 h	1	2	3	4	7	9	10	12	14	21
Cornea	Opacity	A	-	-	1	1	0	-	0	-	-	-	-	-
	Area involved	B	-	-	1	1	0	-	0	-	-	-	-	-
	(AxB) x 5		-	-	5	5	0	-	0	-	-	-	-	-
Iris		C	-	-	1	0	0	-	0	-	-	-	-	-
	C x 5		-	-	5	0	0	-	0	-	-	-	-	-
Conjunctiva	Redness	D	-	-	2	2	1	-	0	-	-	-	-	-
	Chemosis	E	-	-	2	1	1	-	0	-	-	-	-	-
	Discharge	F	-	-	2	1	1	-	0	-	-	-	-	-
	(D+E+F) x 2		-	-	12	8	6	-	0	-	-	-	-	-
Total			-	-	***22***	***13***	***6***	-	***0***	-	-	-	-	-

			Observation period (days)											
Animal No.	5		1 h	4 h	1	2	3	4	7	9	10	12	14	21
Cornea	Opacity	A	-	-	0	0	0	-	0	-	-	-	-	-
	Area involved	B	-	-	0	0	0	-	0	-	-	-	-	-
	(AxB) x 5		-	-	0	0	0	-	0	-	-	-	-	-
Iris		C	-	-	0	0	0	-	0	-	-	-	-	-
	C x 5		-	-	0	0	0	-	0	-	-	-	-	-
Conjunctiva	Redness	D	-	-	1	0	0	-	0	-	-	-	-	-
	Chemosis	E	-	-	0	0	0	-	0	-	-	-	-	-
	Discharge	F	-	-	0	0	0	-	0	-	-	-	-	-
	(D+E+F) x 2		-	-	2	0	0	-	0	-	-	-	-	-
Total			-	-	***2***	***0***	***0***	-	***0***	-	-	-	-	-

			Observation period (days)											
Animal No.	6		1 h	4 h	1	2	3	4	7	9	10	12	14	21
Cornea	Opacity	A	-	-	1	1	1	-	0	-	-	-	-	-
	Area involved	B	-	-	2	4	1	-	0	-	-	-	-	-
	(AxB) x 5		-	-	10	20	5	-	0	-	-	-	-	-
Iris		C	-	-	0	0	0	-	0	-	-	-	-	-
	C x 5		-	-	0	0	0	-	0	-	-	-	-	-
Conjunctiva	Redness	D	-	-	2	2	1	-	0	-	-	-	-	-
	Chemosis	E	-	-	1	1	1	-	1	-	-	-	-	-
	Discharge	F	-	-	3	1	1	-	0	-	-	-	-	-
	(D+E+F) x 2		-	-	12	8	6	-	2	-	-	-	-	-
Total			-	-	***22***	***28***	***11***	-	***2***	-	-	-	-	-

MMAS (Modified Maximum Average Score) (14+13+8+22+2+22) / 6 = 13.5

5. Appreciation of the in vivo data

With the knowledge of the factors influencing the *in vivo* results the "black or white" approach applied by ICCVAM can hardly be defended. Weil and Scala (1971) even concluded that the eye irritation test in rabbits as published by the Federal agencies of the US should not be recommended as standard procedure in any new regulations. However, the Draize eye test has been used practically unchanged until now.

Michael Balls and Julia Fentum (1993), scientists in the field of validation, concluded: *"It is very rare for any allowance to be made for the variability of the animal data, which are thus given a status which they do not deserve. They wrongly become the "true" values which the non-animal tests must struggle to reproduce. Also, insufficient allowance is made for the doubt which must be placed on values which fall within the barrier zones on both sides of category cut-off points. This is particular worrying when Cooper two-by-two way*

plots are used as a basis for establishing the sensitivity, specificity, predictivity and concordance of in vitro test data". Bruner (1996), another scientist in the field of validation, concluded after computer simulations that even if the alternative methods were perfectly reproducibly (if their coefficients of variation were 0), the variability in the Draize scores alone would restrict the Pearson's correlation coefficients to the range 0.65 - 0.80 when the Draize scores are between 0 and 40, which are typical for (mild) irritants.

ICE OECD Test guideline 438

One of the conclusions in the ICCVAM test method evaluation report (2006) was: *"There are sufficient data to support the use of the ICE test method, in appropriate circumstances and with certain limitations, as a screening test to identify compounds as ocular corrosives and severe irritants (i.e., EPA Category I, UN GHS Category 1, EU R41) in a tiered-testing strategy, as part of a weight-of-evidence approach. The identified limitations for this method are based on the false negative and false positive rates that are observed for certain chemical and physical classes. Based on the available database, the false negative rates for alcohols, surfactants and solids range from 33% (1/3) to 50% (1/2), 44% (4/9) to 57% (4/7), and 46% (6/13) to 70% (7/10), depending on the hazard classification system (EC, UN-GHS or EPA) used. Additionally, the false positive rates for alcohols range from 27% (3/11) to 50% (5/10)".* Two of the alcohols in the data base were ethanol and butanol. Both caused severe irritancy in the ICE (and BCOP) but were Category 2 according to the *in vivo* data of ECETOC (1998). As discussed previously the *in vivo* data of Weil and Scala for ethanol (and also for butanol) showed very high inter- and intralaboratory variations, making the ICCVAM conclusion on performance of the ICE with respect to alcohols questionable. Also the surfactant examined in the study of Weil and Scala showed high inter- and intralaboratory variations.

In 2009, on the basis of the ICCVAM evaluation report the ICE and BCOP were adopted as an OECD Test Guideline (TG 438 and TG 437, respectively) for the screening of severe eye irritants. The false negative rates for identifying severe irritants, i.e. compounds identified by the ICE as not severely irritating, were not considered critical since these compounds are still to be tested in the *in vivo* rabbit eye test. In the OECD test guideline, the limitations with respect to the screening of surfactants, alcohols and solids were also mentioned. Specifically it was mentioned that: *"The current validation database did not allow for an adequate evaluation of some chemical or product classes (e.g. formulations). However, investigators could consider using this test method for testing all types of compounds (including formulations), whereby a positive result could be accepted as indicative of an ocular corrosive or severe irritant response. However, positive results obtained with alcohols should be interpreted cautiously due to risk of over-prediction".* The specific mentioning of formulations is remarkable because the ICE is used more frequently for formulations than for pure compounds. Moreover, the Draize eye test also does not make any distinction between the testing of pure compounds and formulations. In general, one should realize that at that time the eye irritation potential of almost all, if not all, compounds had been determined for regulatory purposes in one single type of eye (test), i.e. the rabbit eye (test).

A follow-up evaluation of the usefulness and limitations of alternatives for identifying mild/moderate and non-irritant chemicals was made by NICEATM/ICCVAM, in collaboration with ECVAM and JaCVAM. In early 2011, a proposal for updating the BCOP TG 437 for the identification of chemicals not requiring a classification for eye irritation was submitted to the OECD by means of a Standard Project Submission Form (SPSF). The BCOP database comprised 196 compounds of which 89 were non-irritants, and these data were used to draft a Streamlined Summary Document (SSD). The BCOP was considered appropriate because the percentage of false negatives for non-irritants was 0%. However, the percentage of false positives was 69%. The ICE test was not proposed for such an update, because the review panel maintained the original recommendation to use the ICE only for classification of ocular corrosives and severe irritants. Specific objections against the use of the ICE for chemicals not requiring a classification was the fact that two compounds of the "parallel" dataset (coded TNO-28 and TNO-94) identified as non-irritants by the ICE turned out to be severe irritants in the *in vivo* rabbit eye test. In collaboration with the National Coordinator for the OECD, it was argued by M.K. Prinsen that the review panel had not studied the nature of the effects of these two false negative substances in detail. TNO-94 was an anti-fouling paint for the shipping industry, a specific type of product, which produced reversible irritating eye effects in two out of three rabbits. In the third rabbit an unusual effect occurred, i.e. adherence of the paint to the cornea which was reason to humanely sacrifice the animal on day 1. Anti-fouling paints are designed to be very durable which may explain the findings in this rabbit. Whether or not this peculiar effect is relevant for humans, excluding (anti-fouling) paints from ICE testing would have no major consequences for the applicability of the method for screening of non-irritants in general. TNO-28 caused no corneal or iris effects; only conjunctival effects were observed. The conjunctival effects observed with this compound were below the threshold for classification as an eye irritant. Eye effects had cleared completely in one rabbit after 72 hours and in another rabbit after 7 days. The third rabbit showed increased conjunctival effects at 48 hours after treatment and on day 14 moderate redness and slight swelling of the conjunctivae were still observed. Importantly, a white ocular discharge was also observed which was a sign of secondary infection. The fact that the same effects were observed at 21 days after treatment supported this assumption. One week later, the eye effects in this rabbit had cleared completely.
Overall, the ICE "parallel" data set provided by TNO showed a false negative rate for non-irritants of 6% and a false positive rate of 1%. Therefore, the OECD was asked to reconsider the applicability of the ICE for the purpose of identifying non-irritants. During an OECD expert meeting (6-7 December 2012, Paris), M.K. Prinsen presented the issues of the *in vivo* Draize eye test as discussed previously.

As a result, the limitations of the *in vivo* Draize rabbit eye irritation test and their implications for validation purposes were recognized and summarized in the document (OECD SSD ICE, 2013) as follows:

1. The *in vivo* rabbit eye irritation/corrosion test has no standardized exposure regimen. Therefore, the duration of exposure of the test substance with the rabbit eyes remains unknown and can vary from a few minutes to several hours. In addition, for solids and sticky chemicals it is unclear how much of the compound (solid, paste or liquid) stays in contact with the eye;
2. The limited reproducibility of the Draize rabbit eye test method;
3. The subjectivity in the allocation of the rabbit ocular tissue scores;
4. The type of exposure which does not reflect a potential human accidental exposure;
5. The differences in physiology and sensitivity to tested chemicals between rabbit and human eyes.

The re-evaluation of the ICE ICCVAM dataset showed that individual *in vitro* and *in vivo* classifications of a number of compounds needed further considerations. Discrepancies were found in the final *in vivo* and *in vitro* classifications for a number of compounds which had an impact on the final number of false negative compounds.
After re-evaluation, the ICE test method had an overall accuracy of 82%, a false positive rate of 33%, and a false negative rate of 1% (instead of 6%) for non-irritants, when compared to the *in vivo* data classified according to the UN-GHS. If anti-fouling organic solvent containing paints were excluded from the database, the accuracy of the ICE test method was 83%, the false positive rate 33% and the false negative rate 0%.
In September 2013, the OECD TG 438 for the ICE was officially adopted also to include the identification of non-irritants (in general about 80% of the chemicals tested are non-irritants). This was a huge success for the ICE and for the 3 R's in general, but it still meant that compounds not identified by the ICE as non-irritant or severe irritant have to be tested in the *in vivo* rabbit eye test.

Lessons learned, considerations and recommendations

- For eye irritation with a lack of human data, the combined *in vitro*/*in vivo* (parallel) testing instead of using *in vivo* rabbit eye irritation data from literature provided an ideal setting for i) developing and validating the alternative method, ii) introducing the method to industry and regulatory authorities, iii) getting insight in, and critically address the pros and cons of both the *in vitro* and *in vivo* test system.
- A meaningful validation of an *in vitro* alternative model in the middle range of irritancy (Category 2 classification) cannot be reached with the current *in vivo* rabbit eye irritation data set due to the large variability.
- The selection procedure to select or to accept models suitable as an *in vitro* alternative to eye irritation should be more critical than in the past. Alternatives should have a direct relation to (human) ocular irritancy and be developed on the basis of the mechanistic principles of (human) ocular inflammation, instead of matching Draize eye test results only.
- It should be realized that the existing *in vivo* rabbit eye irritation data does not reflect the inflammatory and recovery processes in humans. Therefore, the data of the Draize eye test are not suitable for the development of *in vitro* models for eye irritation focusing on discrimination between severe, but reversible or severe, but irreversible eye effects.
- The *in vivo* eye irritation test in rabbits should no longer be allowed, and should be completely replaced by alternative methods, for instance the ICE.
- The household and personal care industry should share their strategy to fulfil regulatory demands without the use of the *in vivo* animal test with other (chemical) industries and regulatory authorities.

Recommendations for future validation

- The experimental animal should no longer be regarded as the “Golden Standard”, but reliable, preferably human, *in vivo* data of reference compounds should be available for validation.
- Equipment and protocols for new methods should be standardized.
- An inventory of factors influencing the (regulatory) acceptance of alternatives should be made with all parties involved prior to the practical start of the validation process.

In conclusion

Alternatives to the Draize eye irritation test should preferably make use of *ex vivo* eyes, eye tissue, or eye tissue equivalents in order to measure, both qualitatively and quantitatively parameters similar or identical to those in the clinic. The appropriateness of the presently available alternatives, based on the above mentioned criteria, is as follows:

1. Models using intact isolated eyes with slit-lamp microscope observations and histopathology, e.g. the ICE or IRE;
2. Models using excised corneas with light transmission and fluorescein measurements and histopathology, e.g. the BCOP (Bovine) or PCOP (Porcine);
3. 2D or 3D human corneal epithelium reconstruction models, which have as major disadvantage the lack of the different membranes of the cornea, corneal stroma and corneal endothelium.

References

Balls, M., Fentem, J.H. (1993). On Recognizing and Overcoming Barriers to the Acceptance of Alternative Methods. *Alternative Methods in Toxicology and the Life Sciences* 11, Part 11 A. Mary Ann Liebert, Inc., publishers, New York.

Balls, M., Botham, P.A., Bruner, L.H., Spielmann, H. (1995). The EC/HO international validation study on alternatives to the Draize eye irritation test. *Toxicology in Vitro* 9, 871-929.

Brantom, P.G., Cassidy, S., Eigler, D., Esdaile, D., Fentem, J.H., Liebsch, M. McPherson, J., Pfannenbecker, U., Prinsen, M.K. (2000). Preliminary evaluation of the application of reference standards in the prevalidation and validation of *in vitro* tests for eye irritation. *TNO BIBRA Report No.* 2/3387/00, Volume 2, December 2000.

Bruner, L.H., Carr, G.J., Chamberlain, M., Curren, R.D. (1996). Validation of alternative methods for toxicity testing. *Toxicology in Vitro* 10(4), 479-501.

Burton, A.B.G. (1972). A method for the objective assessment of eye irritation. *Food and cosmetics toxicology* 10, 209-217.

Burton, A.B.G., York M., Lawrence, R.S. (1981).The *in vitro* assessment of severe eye irritants. *Food and Cosmetics Toxicology* 19, 471-480.

Cazelle, E., Eskes,C., Hermann, M., Jones, P., McNamee, P., Prinsen, M., Taylor, H., Wijnands, M.V.W. (2014). Suitability of Histopathology as an Additional Endpoint to the Isolated Chicken Eye Test for Classification of non-extreme pH Detergent and Cleaning Products. *Toxicology in Vitro* (in press; Ref. No. TIV-D-13-00528).

Chan, T., Payor, S., Holden, B.A. (1983). Corneal thickness profiles in rabbits using an ultrasonic pachometer. *Investigative Ophthalmology and Visual Science* 24(10), 1408-1410.

Doughty, M.J., Zaman, M.L. (2000). Human Corneal Thickness and Its Impact on Intraocular Pressure Measures: A Review and Meta-analysis Approach. *Survey of Ophthalmology* 44 (5), 367-408.

Draize, J.H., Woodard, G., Calvery, H.O. (1994). Methods for study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *Journal of Pharmacology and Experimental Therapeutics* 82, 377-390.

EC (1991). Collaborative study on the evaluation of alternative methods to the eye irritation test. *Document XI/632/91-V/E/1/131/91*, part I and part II. Brussels: European Commission.

EC (1993). *Commission Directive 93/21/EEC* of 27 April 1993 adapting to technical progress for the 18th time Council Directive 67/548/EEC on the Approximation of Laws, Regulations and Administrative Provisions Relating to the Classification, Packaging and Labelling of Dangerous Substances. Official Journal of the European Communities L110A, 1-86.

ECETOC (1998). Eye Irritation – Reference Chemicals Data Bank. Technical Report No. 48(2). European Centre for Ecotoxicology and Toxicology of Chemicals, Brussels.

Faber, C., Scherfig, E., Prause, J.U., Sørensen , K.E. (2008). Corneal thickness in pigs measured by ultrasound pachymetry *in vivo*. *Scandinavian Journal of Laboratory Animal Science* 35, 39-43.

Federal Register notice 13513, Volume 70, No. 53, Monday, March 21, 2005: National Toxicology Program; National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Availability of Expert Panel Report on the Evaluation of the Current Validation Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants.

Federal Register notice 43149, Volume 70, No. 142, Tuesday, July 26, 2005: National Toxicology Program (NTP); NTP Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM); Availability of Revised Analyses and Proposed Reference Substances for *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants.

Fowler, W. C., Chang, D.H., Roberts, B.C., Zarovnaya, E.L., Proia, A.D. (2004). A new paradigm for corneal wound healing research: The white leghorn chicken (Gallus gallus domesticus). *Current Eye Research* Vol. 28, No. 4, Pages 241-250.

Grant, W.M. (1986). Corneal epithelial Vacuoles. *Toxicology of the Eye*, 3rd edition, Volume 1, page 9. Charles C. Thomas Publishers, Springfield, Illinois, USA.

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; *Toxicology and applied pharmacology* 55, 501-513.

Griffith, J.F. (1987); The low volume eye irritation test.; *Soap cosmetics chemical specialties* 63(4), 32-63.

ICCVAM Background Review Document (2006); Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Isolated Chicken Eye Test Method. *NIH Publication* No: 06-4513.

Li, H.F., Petroll, W. M., Møller-Pedersen, T., Maurer, J.K., Cavanagh, H.D., Jester, J.V. (1997). Epithelial and corneal thickness measurements by *in vivo* confocal microscopy through focusing (CMTF). *Current Eye Research* 16(3), 214-221.

Lordo, R.A., Feder, P.I., Gettings, S.D. (1999). Comparing and evaluating alternative (*in vitro*) tests on their ability to predict the Draize maximum average score. *Toxicology in vitro* 13(1); 45-72.

Luepke, N. (1985). Hen's egg chorioallantoic membrane test for irritation potential. *Food and Chemical Toxicology* 23, 287-291.

Maurer, J.K., Parker, R.D. and Jester, J.V. (2002). Extent of corneal injury as the mechanistic basis for ocular irritation: key findings and recommendations for the development of alternative assays. *Regulatory Toxicology and Pharmacology* 36, 106-117.

OECD (1981). *OECD Guideline for Testing of Chemicals* No. 405: Acute eye irritation/corrosion, adopted 12 May 1981. Organisation for Economic Co-operation and Development, Paris.

OECD (2001). *OECD Guidelines for the Testing of Chemicals* No. 423: Acute oral toxicity; Acute Toxic Class method, Paris, France.

OECD (2002). O*ECD Guideline for Testing of Chemicals* No. 405: Acute eye irritation/corrosion, adopted 24 April 2002. Organisation for Economic Co-operation and Development, Paris.

OECD (2004). *OECD Guideline for Testing of Chemicals* No. 431: *In Vitro* Skin Corrosion: Human Skin Model Test, adopted 13 April 2004. Organisation for Economic Co-operation and Development, Paris.

OECD (2009). *OECD Guideline for Testing of Chemicals* No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying Ocular Corrosives and Severe Irritants, adopted 7 September 2009. Organisation for Economic Co-operation and Development, Paris.

OECD (2009). *OECD Guideline for Testing of Chemicals* No. 438: Isolated Chicken Eye Test Method for Identifying Ocular Corrosives and Severe Irritants, adopted 7 September 2009. Organisation for Economic Co-operation and Development, Paris.

OECD (2010). *OECD Guidelines for the Testing of Chemicals* No. 429: Skin Sensitisation; Local Lymph Node Assay, Paris, France.

OECD (2010). *OECD Guideline for Testing of Chemicals* No. 439: *In Vitro* Skin Irritation - Reconstructed Human Epidermis Test Method, adopted 22 July 2010. Organisation for Economic Co-operation and Development, Paris.

OECD (2013). *OECD Guideline for Testing of Chemicals* 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, adopted 26 July 2013. Organisation for Economic Co-operation and Development, Paris.

OECD (2013). Draft Streamlined Summary Document (SSD) for the Isolated Chicken Eye (ICE) test method. ENV/JM/TG/RD(2013)19.

Ohno, Y., Kaneko, T., Inoue, T., Morikawa, Y., Yoshida, T., Fuji, A., Masuda, M., Ohno, T., Hayashi, M., Momma, J., Uchiyama, T., Chiba, K., Ikeda, N., Imanashi, Y., Itakagaki, H. (1999). Interlaboratory validation of the *in vitro* eye irritation tests for cosmetic ingredients. (1) Overview of the validation study and draize scores for the evaluation of the tests. *Toxicology in Vitro* 13(1), 73-98.

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 and Chemical Toxicology* 34(4), 291-296.

Prinsen, M.K. (1999). An Evaluation of the OECD Proposal for the Harmonised Classification of Eye Irritants and Corrosives. Report of ECVAM Workshop 34, Eye Irritation Testing: The Way Forward, Appendix 1. *ATLA* 27, 72-77.

Prinsen, M.K. (2005). *in vitro* and *in vivo* data for 94 substances tested in the Isolated Chicken Eye test. Unpublished data provided directly to NICEATM by MK Prinsen, TNO Nutrition and Food Research Institute.

Rama, P., Matuska, S., Paganoni, G., Spinelli, A., De Luca, M., Pellegrini, G. (2010). Limbal Stem-Cell Therapy and Long-Term Corneal Regeneration. *New England Journal of Medicine* 363, 147-155.

Scala, R. A., Springer, J. (1997). IRAG working group 6; Guidelines for the Evaluation of Eye Irritation Alternative Tests: Criteria for Data Submission. *Food and Chemical Toxicology* 35, 13-22.

Schutte, K., Prinsen, M.K., McNamee, P.M., Roggeband, R. (2009).The isolated chicken eye test as a suitable *in vitro* method for determining the eye irritation potential of household cleaning products. *Regulatory Toxicology and Pharmacology* 54(3), 272-281.

Terry Kim, T., Khosla-Gupta, B. A. (2002). Chemical and Thermal Injuries to the Ocular Surface. *Ocular Surface Disease Medical and Surgical Management,* Part III, pp 100-112. Springer-Verlag New York, Inc.

Thoft, R.A., Friend, J. (1977) Biochemical transformation of regenerating ocular surface epithelium. *Investigative Ophthalmology and Visual Science* 16, 14-20.

United Nations-Economic Commission for Europe (UN/ECE) (2007). Globally Harmonised System of Classification and Labelling of Chemicals (GHS). UN, New York and Geneva, 2007. Available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html.

US-FDA (1964). Illustrated guide for grading eye irritation by hazardous substances. Washington, D.C. 20204, USA.

Wagoner, M.D. (1997). Chemical Injuries of the Eye: Current Concepts Pathophysiology and Therapy. *Survey of Ophthalmology* 41(4), 367-408.

Weil, C.S., Scala, R.A. (1971). Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests. T*oxicology and applied pharmacology* 19, 276-360.

Summary

10

Introduction

On 2 November 1944 a manuscript, entitled "Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes" was received for publication by the Journal of Pharmacology and Experimental Therapy. The authors of this article were John H. Draize, Geoffrey Woodard and Herbert O. Calvery from the Division of Pharmacology, Food and Drug Administration, Federal Security Agency, Washington, D.C., USA. It is more than likely that the authors never expected the kind of impact this publication would have on animal experimentation worldwide. Almost seventy years later the name Draize is still inextricably attached to one of the most disputed toxicity tests commonly used to determine acute toxicity, i.e. the Draize eye irritation test. To establish the potential risk of exposure of the eyes to compounds, the Food and Drug Administration of the United States (US-FDA) adopted the Draize eye irritation test using rabbits already in 1961. Internationally, the Organisation for Economic Co-operation and Development (OECD) published their first guideline on eye irritation in 1981, which was subsequently adopted by the European Union in 1984. In the early eighties, the controversial character of this type of animal testing became known to the general public and the development of alternative non-animal tests to replace the Draize eye test started. The publication of Russell and Burch in 1959 entitled: "The principles of humane experimental technique" formed the basis of most initiatives relating to the use and development of alternatives for animal experiments. In their publication they postulated the famous and often cited three R's: Reduction, Refinement and Replacement of animal experiments. The 3 R initiatives concerning the Draize eye test mainly involved reduction of the number of animals from six to three per test. Certain aspects of the Draize eye test causing considerable pain and discomfort to the animal were dealt with only at a much later stage, i.e. in 2002 the reduction of the time for a wash-out of the test substance from 24 hours to 1 hour after instillation, and in 2012 the use of systemic pain relief and topical sedation.

The first attempts to validate alternative tests for eye irritation were considered to be relatively simple by comparing *in vitro* and *in vivo* irritation index scores. What was expected to be a process of several years, however, turned out to be a decades spanning process still not fully completed. For a large part, this can be attributed to the nature of the *in vivo* test in rabbits, which is more complicated and compromised than originally believed.

In the early nineteen-eighties, the TNO-CIVO Toxicology and Nutrition Institute in Zeist, the Netherlands, introduced an alternative test for the Draize eye irritation test. This alternative, using isolated eyes, was based on a method developed by A.B.G. Burton from Unilever. This thesis describes the development, performance, validation and acceptance of the *in vitro* isolated eye test which focuses on measuring the damage of compounds to the cornea.

Chapter 2 describes the results of the first in-house validation of the isolated rabbit eye test (IRE) at TNO-CIVO. Substances, already tested in the *in vivo* eye irritation test at the request of various industries and representing the average supply of substances in contract research, were tested in the IRE with a modified superfusion apparatus and eye-clamps. The IRE results showed an overall correlation of 82% with the *in vivo* results, using a general classification scheme for the grading of *in vivo* and *in vitro* eye irritation (i.e. not, slight, moderate or severe irritant). Four compounds, all hydrophobic, were overpredicted by the IRE. A general observation was that hydrophobic compounds stayed in contact with the cornea (eye) of the rabbit for a relatively short period of time, mainly because its third eye-lid acted as a wiper. The presence of a third eye-lid is an example of a condition very specific to the *in vivo* rabbit eye test influencing the exposure to (hydrophobic) substances. This caused a problem with respect to the validation of alternative methods since alternative methods are not able to mimic such a condition. Moreover, it is a non-existing condition in man.

Although rabbits are available as eye-donor for the isolated eye test in contract research organisations executing routine eye- and skin irritation studies, the dependency on laboratory animals was considered a serious shortcoming. **Chapter 3** describes the use of eyes of animals from slaughter-houses as a source for eyes. Slaughterhouse animals, such as the cow, pig or chicken were considered as eye-donors. Although the pig eye was considered better comparable to the human eye, the chicken eye proved to be the most suitable candidate. The isolated eye test had to produce matching results with the Draize rabbit eye test in order to be accepted as an alternative. Hence, an eye (cornea) that matched closest to the rabbit eye and not to that of the human eye was needed. The chicken had such an eye because its corneal thickness is similar to the rabbit cornea. The epithelium of the cornea is the first barrier against (chemical) insult and as such its thickness (number of layers of epithelial cells) is of importance. The baseline corneal thickness of the cow and pig was already that high that measurement of increased thickness to its full extend after treatment with moderate to severe irritants was not possible. The suitability and sensitivity of chicken eyes was evaluated by testing 21 reference chemicals, which were selected to be representative of currently used industrial chemicals of different chemical classes, ranging from non-irritant to severe irritant. Furthermore, a criteria and scoring system of the *in vitro* corneal effects together with a Prediction Model (PM) for matching the EC scheme for classification and labelling of compounds were introduced for the isolated chicken eye (ICE) test. On the basis of the results it was concluded that:
i) although the ICE does not assess conjunctival damage, its sensitivity to predict ocular damage is not reduced, ii) the ICE correctly predicted the EC classifications of the 21 reference chemicals and iii) the ICE fitted in the previously updated EC and OECD guidelines regarding acute eye irritation/corrosion including recommendations to use alternatives for the prescreening or positive identification of strong eye irritants. Thereafter, the ICE was incorporated as a prescreen in the routine *in vivo* assessment of eye irritation testing in the frame of contract research at TNO.

In **Chapter 4** the results are presented of the "parallel" (first *in vitro* followed by *in vivo*) testing of 44 compounds at TNO. These compounds were considered to be a relevant cross-section of compounds routinely produced by the (chemical) industry. The *in vivo* scores were compared to the critical scores of the ICE test, namely the scores for corneal swelling, corneal opacity, fluorescein retention of damaged corneal epithelium and an index score (combination of the three ICE parameters). The overall correlations found for the *in vivo* scores with the ICE *in vitro* scores were 0.90 (index score), 0.91 (corneal swelling), 0.86 (corneal opacity) and 0.82 (fluorescein retention). The correlation between the *in vivo* conjunctival scores and the ICE scores were 0.92 (index score), 0.92 (corneal swelling), 0.93 (corneal opacity) and 0.86 (fluorescein retention). These correlations substantiated the conclusion made earlier by Burton in 1972 that a relationship exists between the *in vivo* conjunctival damage and the corneal scores of the isolated eye test. Overall, it was concluded that the ICE provided a practical prescreen for the Draize rabbit eye test and that only mild to moderate irritants in the ICE, generally showing the highest sensitivity to inter- and intra-laboratory variability, should be confirmed in the rabbit eye test. Eighteen years later, the OECD adopted this conclusion. The "parallel" *in vitro* and *in vivo* eye irritation testing was continued with another 50 compounds and the results were submitted to organizations dedicated to the validation of alternative non-animal tests, such as ECVAM (European Centre for the Validation of Alternative Methods) and ICCVAM (Interagency Center for the Evaluation of Alternative Toxicological Methods) of the USA.

At TNO the ICE test was also used as a stand-alone test, especially for the household and personal care industry, which increasingly adopted non-animal testing strategies. The Procter & Gamble Company was one of these companies that used the ICE test for their eye irritation safety program, and **Chapter 5** describes the application of the ICE test to their domain of household cleaning products. In the past, P&G had heavily invested in the Low Volume Eye Test (LVET), a modification of the Draize eye test. Because of the rather extreme exposure conditions of the standard Draize eye test, their products often needed labelling as severe eye irritants. The LVET used only one tenth of the dose volume of the test material (0.01 instead of 0.1 mL/gram) administered directly onto the cornea instead of instilling it in the lower conjunctival sac of the rabbit eye. The results of the LVET correlated well with their database on human responses to accidental exposure and to results of clinical studies. Although less burdening for the animals, the LVET was still an animal test and in view of their non-animal testing policy the suitability of the ICE for their products was explored. Twenty products representing cleaners, dishwashing liquids or powder, powder detergents and raw materials of these products were tested at TNO in the ICE employing different dosing volumes to examine the correlation with the LVET. Furthermore, histopathology of the treated corneas sampled at the end of the observation period was included. With histopathology the correlation between the depth of injury in the cornea and possible recovery of eye lesions could be studied. Since measurement of recovery is not possible in the ICE (and in all other alternatives), it could be an important

improvement of the method. The outcome of this study was that, in general, the results of the ICE using its standard dosing volume of 30 µl or 30 mg were in line with or more conservative than the LVET in terms of classification. The level of overprediction found in the ICE was expected and considered acceptable since the ICE was developed for prediction of the Draize eye test and not the LVET. For P&G the ICE test was considered useful for several purposes, such as 1) EU/GHS classification and labelling of powder and liquid household cleaning products, 2) screening of candidate formulations, and 3) weight-of-evidence approach by determining the profile of new cleaning product formulations against benchmark products. A definite conclusion on the usefulness of histopathology in the ICE could not be made, but the data showed that assessment of the histopathological lesions in the various parts of the cornea was possible, enabling the assessment of the depth of injury in the cornea.

Chapter 6 deals with investigations by TNO in the search for additional parameters that could be helpful to discriminate between the different irritancy levels in the ICE. For that purpose, histopathology of the cornea using different staining techniques were examined. Originally, Haematoxylin & Eosin (H&E) was used for staining the microscopic slides of the cornea. Later on H&E staining was replaced by Periodic Acid-Schiff (PAS) staining, which provided a better discrimination of the different layers of the cornea, i.e. the epithelium, stroma and endothelium and other structures, such as the basement membrane and the Bowman's membrane (between epithelium and stroma) and the Descemet's membrane (between stroma and endothelium). The integrity of these membranes was considered to play a role in the injury and recovery process of the cornea, and visibility of these membranes by microscope was considered important for an adequate histopathological assessment by the pathologist. Therefore, staining methods such as Trichrome, AZAN (Azocarmine & aniline) and EVG (Elastic Van Gieson), specifically targeting collagen-rich membranes, were tested on corneas treated with compounds representing a non-irritant, irritant and severe irritant (corrosive). PAS appeared clearly superior with respect to visibility of the membranes and the quality of the morphology of the various corneal structures. The histopathological examinations also showed that after severe corrosive damage to the cornea by sodium hydroxide, the basement and the Bowman's membranes appeared undamaged while lesions were seen in the underlying stroma. This observation led to the conclusion that depth of injury is not the only factor determining the seriousness of corneal injury. In the clinic, emphasis is put on corneal opacity and corneal stem cell survival after (chemical) injury in order to evaluate its severity. With severe stem cell damage, recovery of the corneal damage by re-epithelialization of migrating stem cells from the corneal limbus is not possible. In that case, the recovery process will result in complete, irreversible conjunctivalization of the ocular surface. The possibility of screening the viability of stem cells after chemical injury could, therefore, be of value. However, explorations to stain stem cells of the chicken cornea by p63 immunostaining by the histopathology section of TNO appeared unsuccessful. Although the assessment of reversibility or irreversibility in the ICE is undoubtedly of value, especially

for household and personal care companies, the EU/GHS classification and labelling of severe eye irritants does not make a distinction between these two categories; they are all classified as Category 1: "irreversible effects on the eye/serious damage to the eye".

The need for accepted alternative methods led to international validation studies involving several promising alternatives. One of the most comprehensive international validation was held by the European Commission (EC) and the British Home Office (HO) in the mid nineteen-nineties. Sixty chemicals were tested by nine alternative methods, including the ICE, and each performed by four independent labs. With a maximum overall correlation that ranged from 0.34 to 0.55, the outcome was very disappointing which was thought to be due to the use of the Modified Maximum Average Score (MMAS) as the sole parameter for evaluation of *in vivo* eye irritation (**Chapter 7**). The MMAS, which can range from 0-110, is an average of the maximum individual tissue scores of the animals recorded for a compound. A compound was classified as a non-irritant with a MMAS of 0-25, irritant with a MMAS of 25-59, or severe irritant with a MMAS >59. Because the MMAS is not used for regulatory classification, the impact of the EC and proposed OECD harmonized classification system (later on to be adopted as the EU-CLP and UN/GHS classification systems) was investigated by Menk Prinsen. MMAS cut-offs of 25 and 59 belonging to, respectively, irritant and severe irritant, appeared not appropriate for classification according to these two regulatory systems. Eight compounds with an MMAS lower than 59 were in fact severe irritants and 4 compounds with an MMAS higher than 59 and 3 compounds with an MMAS lower than 25 were irritants. It was, therefore, recommended that future validation of alternatives would benefit from the use of classifications based on the proposed OECD harmonized classification system.

In 2004, a new initiative was undertaken to (re-)validate alternative methods which were considered the most promising in the EC/HO study, especially for the screening of severe irritants. It was an initiative of ICCVAM in collaboration with ECVAM and the methods selected were the ICE, the Hen's Test - Chorioallantois Membrane (HET-CAM), the Isolated Rabbit Eye test (IRE) and the Bovine Corneal Opacity and (fluorescein) Penetration test (BCOP). An independent Expert Panel was established for each alternative method to determine the validation status of these methods. In general, the idea was that the poor correlation of the alternatives with the Draize Eye Test was due to shortcomings in the practical performance of the *in vitro* methods which should be improved, rather than also critically addressing the shortcomings of the Draize eye test and its consequences for the validation of *in vitro* alternatives. Following a meeting of the panel experts and experts presenting the alternative methods, a discussion paper on this subject was written by Menk Prinsen, entitled: "The Draize Eye Test and *in vitro* alternatives; a left-handed marriage?" (**Chapter 8**) in which the issues of the *in vivo* test in relation to the development and validation of alternatives were addressed. In general, the *in vivo* results were used as the "Golden Standard" for all comparisons with the alternative tests. Although there

was general agreement that the *in vivo* test in rabbits is far from perfect, the implications of the inconsistencies of the *in vivo* test for the validation of the alternatives were never taken into account. It was generally believed that the *in vivo* rabbit eye test produced variable results due to the fact that different labs and observers were involved and that data were used that had been obtained over a very long period of time. However, several other important issues played a crucial role in the outcome of the *in vivo* test. Instillation of the compound in the conjunctival cul-de-sac of the eye and the almost immediately release of the rabbit thereafter resulted in an exposure, which can be anything from minutes (liquids) to 24 hours (solids). With solids, no washing out of remnants from the conjunctival sac was allowed before 24 hours after treatment. These undefined exposure conditions are in contrast to the basic principles of toxicity testing that advocates a well-defined and controlled exposure to substances. Moreover, this kind of exposure condition by placing a large amount of compound in a retracted eye-lid will hardly occur in humans. Instead of questioning the *in vivo* exposure conditions with solids, ICCVAM considered this to be a deficiency of the *in vitro* method and recommended that the test method needed to be optimized with respect to the exposure conditions for solids. Treatment of the eye post-exposure by the observer can also play an important role. Enclosure of solid materials up to 24 hours in the conjunctival cul-de-sac can result in a complete closure of the eye-lids by the abundant production of colloidal discharge which often forms a sealing crust. If the animal (treated eye) is not receiving special care of the eye an otherwise irritating compound can easily become a severe one. Also these kind of circumstances are not representative for the human situation in case of ocular damage. Another phenomenon that occasionally occurred in the Draize eye test and which was not taken into account when using *in vivo* data, is the development of a secondary eye infection. Because of the inflammation caused by a compound, the eye is more vulnerable to microbiological infection, causing initial mild to moderate effects during the first days after exposure developing into more severe and prolonged effects during the 21 day observation period. It was concluded that after almost 18 years of validation without any real success in replacing the Draize eye test, that further attempts would be futile, if ''old'' *in vivo* data or new data generated by the current OECD guideline were still used for validation. A multi-way approach was recommended by i) implementation in the guidelines (legislation) of the most current *in vitro* methods in the testing strategy for screening of severe irritants, ii) re-evaluation of the Draize Eye Test by standardizing the exposure conditions for liquids and solids and simultaneously iii) a more mechanistically-based development and optimization of *in vitro* methods as an ongoing process.

Chapter 9 of this thesis presents a general discussion of the results obtained. In addition, lessons learned and recommendations for future validation are presented. Overall, the results presented in this thesis showed that alternatives to the Draize eye irritation test should preferably make use of ex vivo eyes, eye tissue (isolated cornea), or eye tissue equivalents (reconstructed corneal epithelium) in order to measure, both qualitatively and quantitatively parameters similar or identical to those in the clinic.

Models using intact isolated eyes with slit-lamp microscope observations and histopathology, like the ICE or IRE are considered the most appropriate for that purpose.

Samenvatting

11

Met de sterke opkomst van de chemische industrie in de 20ste eeuw werd duidelijk dat de omstandigheden op de werkplek nadelige effecten op veiligheid en gezondheid van de mens konden hebben. Kortdurende en langdurige blootstelling aan verschillende chemische stoffen veroorzaakten een scala aan ziektes, variërend van milde niet levensbedreigende schade tot levensbedreigende ziektes als kanker. Naarmate werknemers en consumenten mondiger werden, nam de noodzaak toe om mogelijke blootstelling aan chemische stoffen te bepalen, de mensen vooraf te waarschuwen voor de gevaren en ze hiertegen te beschermen. Een van de vele risico's bij het omgaan met (gevaarlijke) stoffen is het in de ogen krijgen van het product.
Al in 1961 schreef de Amerikaanse toelatingsautoriteit de "Food and Drug Administration" (FDA) een test voor met konijnen om stoffen te testen op hun oog irriterende eigenschappen (de zogenaamde Draize test). Het controversiële karakter van de test werd pas bij het grote publiek bekend nadat Henry Spira, lid en oprichter van een dierenrechtenorganisatie, een paginagrote advertentie in de New York Times plaatste met de kop "Hoeveel konijnen maakt Revlon blind omwille van schoonheid". Binnen een jaar werd geld gedoneerd door diverse cosmetische bedrijven en werd een centrum voor het ontwikkelen van alternatieven voor proefdiertesten opgezet; CAAT (Centre for Alternatives to Animal Testing). De ontwikkeling en validatie van alternatieven voor de oogirritatie test in konijnen, die als een relatief simpele test werd beschouwd, bleek een taai en langdurig proces van vele jaren te worden. Sterker nog, de test is nu na meer dan 30 jaar nog steeds niet volledig vervangen. Dit proefschrift beschrijft de ontwikkeling, optimalisatie, validatie (binnen TNO) en toepassing van een alternatieve test met geïsoleerde ogen van kippen (de Isolated Chicken Eye, kortweg de ICE test) en in bredere zin de internationale validatie en acceptatie van de ICE test door overheidsinstanties. Een groot gedeelte van de discussie betreft ook de uitvoering van de Draize test zelf, omdat de manier waarop de resultaten van deze test gebruikt worden het grootste obstakel bleek te zijn bij de validatie van de alternatieven.

Het oog is één van onze belangrijkste zintuigen om contact te hebben met onze omgeving. We hebben er maar twee en zijn als zodanig kwetsbaar voor beschadiging, veroorzaakt door opzettelijke of onopzettelijke blootstelling aan stoffen. Dit proefschrift behandelt een belangrijk onderdeel van het oog, namelijk het hoornvlies (de cornea) die de buitenste laag vormt en "het venster naar de wereld" is. De cornea bestaat uit diverse lagen waarin geen bloedvaten lopen. Dit laatste is belangrijk voor de geïsoleerde ogen test, omdat er geen directe bloedtoevoer nodig is om de cornea functioneel te houden gedurende langere tijd. De diverse lagen van buiten naar binnen zijn:
1) epithelium, 2) stroma en 3) endothelium, met tussen 1) en 2) het basaal membraan en Bowman's membraan en tussen 2) en 3) het Descemet's membraan (zie figuur 2, Introductie). De integriteit van de lagen en membranen zorgen voor een heldere cornea waarbij het epithelium een eerste barrière vormt met als doel om materialen, zoals stof, water en bacteriën buiten te houden. Tevens vormt het epithelium een glad oppervlak dat zuurstof en voeding uit traanvocht kan absorberen dat vervolgens verder door de cornea wordt getransporteerd. Tevens zitten in het epithelium duizenden

zenuwuiteinden die de cornea zeer gevoelig maken voor pijn.
Al in 1944 publiceerde de Amerikaan John Draize een artikel met daarin een beschrijving van onderzoeksmethoden voor het vaststellen van toxiciteit en irritatie van stoffen die op huid en slijmvliezen worden toegediend, waaronder de oogirritatie test en de LD50 (Lethale Dosis) test. Ruim zeventig jaar later is zijn naam nog steeds onlosmakelijk verbonden met deze zeer omstreden testen. Al bij het eerste gebruik als regulatoire test werd erkend dat het subjectief scoren van de oogeffecten problematisch was.
Om het scoren te standaardiseren werd in 1964 door de FDA een gids met foto's uitgebracht die de onderzoeker kon gebruiken bij het scoren van effecten, zoals opaciteit (troebeling) van de cornea, roodheid en zwelling van de oogleden (zie figuur 3, Introductie). Internationaal adopteerde de OESO (Organisatie voor Economische Samenwerking en Ontwikkeling) de Draize oogirritatie test in 1981, gevolgd door de Europese Unie in 1984. Sindsdien zijn de richtlijnen verschillende keren aangepast, maar de praktische uitvoering van de test bleef onveranderd. De aanpassingen betroffen vooral het aantal dieren per test en maatregelen om onnodig leed te voorkomen, zoals uitsluiting van stoffen met extreme pH's, huid corrosieve stoffen en een gefaseerde start (eerst met één konijn en afhankelijk van de ernst van de effecten door tot drie konijnen). De uitvoering van de test is vrij simpel. De oogleden van het oog worden opengehouden en 0.1 milliliter of 0.1 gram van de proefstof wordt in het onderste, uitgetrokken oogzakje gedeponeerd. Vervolgens worden de oogleden gedurende één seconde gesloten waarna het dier losgelaten wordt. Het andere oog blijft onbehandeld en dient ter controle. Het konijn wordt teruggezet in zijn kooi en is vrij om de proefstof uit zijn oog te verwijderen. De ogen worden 1, 24, 48 en 72 uur na blootstelling beoordeeld volgens een score schaal (zie tabel 2, Introductie). De effecten worden wekelijks, tot 3 weken na blootstelling, vervolgd om vast te stellen of deze effecten reversibel (herstellen) of irreversibel zijn. Hoewel er verschillende internationale richtlijnen waren op het gebied van oogirritatie was de uitvoering praktisch identiek. De systemen voor de classificatie en het labelen van de stoffen op basis van de waargenomen oogeffecten waren echter zeer verschillend. Dit was zeer nadelig voor het valideren van de alternatieve methoden. Daarom was het implementeren van een geharmoniseerd classificatie systeem in 2007 door de UN (United Nations) een aanzienlijke verbetering.

Aan de basis van het vaak geciteerde 3 V principe, Vermindering, Verfijning en Vervanging van proefdiergebruik, staat de publicatie van Russell en Burch uit 1959, getiteld: "The Principles of Humane Experimental Technique". De initiatieven met betrekking tot de Draize oogirritatie test betroffen vooral het verminderen van het aantal dieren per test van zes naar drie en mogelijke vervanging door proefdiervrije alternatieven. Pas in een veel later stadium werden maatregelen genomen om bepaalde onderdelen van de test, die zeer veel leed kunnen veroorzaken, te verbeteren.
Het werd bijvoorbeeld pas in 2002 toegestaan om proefstof (vooral vaste stoffen), die tot 24 uur na blootstelling nog in het oogzakje aanwezig kon zijn, na 1 uur i.p.v. 24 uur te verwijderen. In 2012 werd de OECD richtlijn aangepast met duidelijke aanwijzingen voor lokale en systemische pijnverlichting voor de konijnen. Begin tachtiger jaren van

de vorige eeuw introduceerde Herman Koëter, lid van de werkgroep "Kritische Evaluatie Toxiciteitstesten" van de Nederlandse Vereniging van Toxicologen, de geïsoleerde ogen test op TNO-CIVO in Zeist. Verschillende alternatieven voor oogirritatie, variërend van bevruchte kippeneieren, cellijn toxiciteitstesten tot sperma motiliteit, waren al gepubliceerd, maar de geïsoleerde ogen test werd gekozen omdat bij oogirritatie de cornea het belangrijkste doelorgaan is.

In 1981 publiceerde A.B.G. Burton van Unilever een methode met geïsoleerde konijnenogen om ernstig irriterende stoffen te identificeren. Hij had al in 1972 ontdekt dat het meten van de dikte van de cornea na blootstelling bij konijnen een objectieve en nauwkeurige vaststelling van oogirritatie opleverde. Uit wetenschappelijk oogpunt is het gebruik van geïsoleerde ogen zeer aantrekkelijk, immers er wordt een oog (*ex vivo*) voor een oog (*in vivo*) gebruikt en daarbij worden effecten (parameters) gemeten die direct vertaald kunnen worden naar die van het dier, maar ook naar die van de mens. Met financiële steun van de Dierenbescherming en de Stichting Schoonheid Zonder Wreedheid werd de benodigde apparatuur aangeschaft en gebouwd. Als eerste werd de methode geëvalueerd door de stoffen die Burton getest had, ook te onderzoeken in de TNO opstelling. In **Hoofdstuk 2** wordt de implementatie en validatie beschreven van de Isolated Rabbit Eye (IRE) gedurende de periode 1982-1985. Verbeteringen van het superfusie apparaat en ooghouders werden aangebracht en met de opstelling werden vervolgens 34 stoffen getest, die net daarvoor waren onderzocht in de Draize test. Er waren geen extra dieren voor de IRE test nodig, omdat konijnen gebruikt werden die al waren ingezet voor huid of oogirritatie (controle oog). Gebruikmakend van een algemeen classificatie schema (onderscheid in niet, licht, matig en ernstig irriterend) werd in 82% van de gevallen de juiste correlatie gevonden. Slechts in vier gevallen was er een overschatting (overprediction), d.w.z. classificatie als licht of matig irriterend (*in vitro*) in plaats van niet irriterend (*in vivo*). Opvallend was dat dit hydrofobe stoffen betroffen die bij konijnen snel uit het oog werden verwijderd doordat deze stoffen slecht mengden met de aanwezige traanfilm en door het derde ooglid dat als een soort wisser fungeerde. Dit laatste is een voorbeeld van een conditie in de Draize test die overduidelijk de blootstelling beïnvloedt, terwijl dit geen rol speelt bij humane blootstelling. Op basis van deze resultaten werd geconcludeerd dat de IRE een gevoelig en bruikbaar testsysteem was voor het identificeren van oog irriterende stoffen. Niet irriterende stoffen in de IRE, waarbij frequent oogcontact werd verwacht, zouden dan alsnog in de Draize test moeten worden getest.

Hoofdstuk 3 beschrijft een belangrijke mijlpaal in de ontwikkeling van de geïsoleerde ogen test, namelijk het gebruik van slachtdieren als oog donor. Hoewel in die tijd de meeste onderzoekslaboratoria over voldoende, al eerder gebruikte, konijnen beschikten om de test uit voeren, werd het hergebruik van proefdieren nog steeds als tekortkoming van de methode gezien. Daarom werd het gebruik van slachtdieren zoals het rund, het varken en de kip als donoren van ogen. Als eerste werd nagegaan of het mogelijk was om ogen te verzamelen in het slachthuis. Daarna werd de praktische kant van het gebruik van verschillende typen ogen in de testopstelling getoetst. Al snel

viel het rund als donor af - te onregelmatige aanvoer, te verschillende achtergrond van de dieren en teveel beschadigde ogen. Bovendien bleek de cornea (circa 1000 µm) extreem dikker t.o.v. de konijnen cornea (circa 400 µm). Het verzamelen van varkensogen op het abattoir was moeilijk, maar mogelijk. Toch viel ook het varkensoog af, ondanks dat dit oog vooraf als een goede kandidaat werd beschouwd vanwege de overeenkomsten in fysiologie met die van de mens, omdat ook deze cornea te dik bleek (circa 600 µm). Doordat de basisdikte van de cornea al in het top bereik van de afleesschaal van het diktemeetapparaat van de opstelling lag, zouden de waarden bij matig tot ernstig irriterende stoffen mogelijk niet meer volledig bepaald kunnen worden. Andere mogelijkheden voor de meting in dat bereik waren ook niet voorhanden.
De geïsoleerde ogen test moest vergelijkbare resultaten leveren t.o.v. de *in vivo* test in konijnen om te worden geaccepteerd als alternatief. Daarom was een oog (cornea) nodig die het beste vergelijkbaar was met die van het konijn en niet met die van de mens. De kip heeft een dergelijk oog, zowel de afmetingen als de dikte van de cornea (circa 400 µm) kwamen overeen met die van het konijn.
Het verkrijgen van kippenogen bleek relatief eenvoudig door de koppen te verzamelen nadat de kippen waren gedood aan de slachtlijn. De koppen werden vervoerd naar het laboratorium in plastic dozen om daar de ogen uit de koppen te ontleden en in het superfusie apparaat te plaatsen. Dit alles gebeurde binnen 2 uur na het verzamelen van de koppen aan de slachtlijn. De ogen die ingezet werden moesten onbeschadigde corneas hebben, d.w.z. geen of zeer geringe troebeling van de cornea, geen of zeer geringe opname van fluoresceïne door beschadigde of dode epitheel cellen en een dikte die in de normale range lag. De geschiktheid van het kippenoog werd vastgesteld met 21 referentiestoffen (van niet irriteren tot ernstig irriterend) die al eerder in een EC (pilot) validatie studie van de IRE waren getest. Een belangrijke ontwikkeling daarbij was het opstellen van de criteria voor het scoren van de effecten en een voorspellingsschema (Prediction Model = PM) voor classificatie op basis van het toen geldende EC systeem voor labeling en classificatie van stoffen (zie tabel 3 van **Hoofdstuk 2**). Omdat de effecten die in de ICE gemeten werden een directe relatie hadden met de effecten die in de Draize test optraden (bijvoorbeeld ten aanzien van troebeling), kon de PM vooraf voor het grootste gedeelte theoretisch vastgesteld worden op basis van de reikwijdte van de response (troebeling, zwelling en fluoresceïne opname). Dit in tegenstelling tot de meeste andere alternatieven die effecten maten die geen directe relatie met oogschade hadden. Deze alternatieven konden alleen na het testen van meerdere stoffen achteraf door middel van een wiskundige vergelijking het *in vitro* effect(en) vertalen naar een *in vivo* effect. Bijvoorbeeld, in de HET-CAM test met bevruchte kippeneieren werd het chorioallantois membraan, rijk aan bloedvaten, blootgesteld aan de stof. Daarna werd de tijd bepaald van het oplossen van de bloedvaten en/of ontstaan van bloedingen na de start van de blootstelling. Vervolgens werd het aantal seconden via een wiskundige vergelijking omgezet naar een getal dat vergeleken kon worden met een *in vivo* irritatie score. Pas na het testen van een groot aantal stoffen kon een dergelijke PM worden bepaald en vaak bleek dat de PM moest worden bijge-

steld als meer stoffen werden getest. De ICE gebruikte een andere benadering, niet door de ICE data te vertalen naar één *in vitro* score en die te vergelijken met de *in vivo* score, maar door de vooraf bekende reikwijdte van elke parameter in te delen in logische categorieën, namelijk in geen effect (categorie I), licht effect (categorie II), matig effect (categorie III) of ernstig effect (categorie IV).
Dus bij het testen van een stof werd één categorie per parameter vastgesteld, namelijk voor troebeling, voor zwelling en voor fluoresceïne opname. De vaststelling van de eind classificatie in niet irriterend, (licht/ matig) irriterend of ernstig irriterend werd bepaald door de combinatie van deze 3 categorieën. Opnieuw was dat een puur theoretische aangelegenheid op basis van de mogelijke combinaties. Bijvoorbeeld in het niet irriterende gebied hoort de combinatie van de categorieën I/I/I en in het ernstig irriterende gebied de combinatie van de categorieën IV/IV/IV. Alle mogelijke combinaties en bijbehorende classificaties zijn terug te vinden in tabel 1 van **Hoofdstuk 9**.
Pas in 2007 was het nodig het classificatie systeem van de ICE aan te passen omdat er toen een internationaal schema, het United Nations - Globally Harmonized System (UN-GHS), in werking trad. Vervolgens moest het schema van de de ICE in 2013 nog een keer aangepast worden om geaccepteerd te worden als OECD richtlijn voor het identificeren van niet-irriterende stoffen.

In toenemende mate werd de ICE ook gebruikt als een op zichzelf staande test zonder vervolg door een *in vivo* test. Vooral bedrijven die een veiligheidsprogramma hanteerden waarin geen dier testen meer voorkwamen, waren geïnteresseerd in de ICE. Eén van de grootste producenten op het gebied van huishoudelijke en verzorgingsproducten, de Procter & Gamble Company (P&G), adopteerde de ICE voor het screenen op oogirritatie. Dit was een enorme stimulans voor de verdere toepassing en ontwikkeling van de ICE. Voor P&G was het testen van hun zeepmiddelen in de Draize test een groot probleem. Met de blootstellingscondities van de Draize test werden vooral zeeppoeders als ernstig irriterend geïdentificeerd. Daarom hadden ze in het verleden sterk ingezet op een modificatie van de Draize test, de “Low Volume Eye Test (LVET)” waarbij het konijnenoog niet aan 0.1 mL of 0.1 gram, maar aan 0.01 mL of 0.01 gram werd blootgesteld. Daarbij werden deze sterk gereduceerde hoeveelheden niet in het oogzakje maar direct op de cornea toegediend. Deze condities, die veel realistischer waren t.o.v. mogelijke humane blootstelling, gaven resultaten die meer in lijn lagen met geregistreerde schadegevallen en klinische studies. Toch bleef de LVET een dierstudie, weliswaar met minder ongerief voor de dieren, maar niet langer passend in hun dierproefvrije strategie. Daarom besloot P&G de mogelijkheden van de ICE te laten onderzoeken bij TNO. In **Hoofdstuk 5** worden de resultaten van de ICE met een specifieke groep van hun producten, namelijk de huishoudelijke schoonmaakmiddelen beschreven. Naast de standaard hoeveelheid van 30 microliter of 30 milligram in de ICE, die correleren met de hoeveelheden in de Draize test, werden ook 10 en 3 microliter of 10 en 3 milligram onderzocht om vast te stellen of de resultaten beter correleerden met de hoeveelheden gebruikt in de LVET. Ook werd voor het eerst histopathologisch onderzoek van de cornea toegevoegd aan de ICE. Met dit extra onderzoek zou mogelijk

de diepte van de schade in de cornea (Depth-of-Injury) kunnen worden vastgesteld. Microscopisch onderzoek was ook al toegepast in de LVET en daaruit bleek een correlatie tussen de diepte van de schade en mogelijk herstel (reversibiliteit) van de oogeffecten. Aangezien de ICE geen herstel van oogschade kon vaststellen, zou het microscopisch vaststellen van de diepte van de effecten kunnen bijdragen aan het beargumenteren van mogelijk herstel van de vastgestelde oogschade. De uitkomst was dat de resultaten met de standaard ICE hoeveelheid van 30 microliter of milligram in lijn of conservatiever waren dan die van de LVET. De mate van overschatting werd acceptabel geacht omdat de ICE ontwikkeld was om de Draize test, en niet de LVET, te voorspellen. P&G achtte de ICE geschikt voor meerdere van hun doeleinden, zoals bijvoorbeeld de EU/GHS classificatie van hun schoonmaakproducten. Dit hield in het vooraf testen van hun kandidaat producten en het toepassen van een benadering met meervoudige bewijsvoering ("weight-of-evidence" benadering) door middel van vergelijking met bestaande producten. Een duidelijke uitspraak over de toegevoegde waarde van histopathologie in de ICE kon niet worden gegeven, maar het bleek wel mogelijk om effecten in de diverse delen van de cornea vast te stellen waardoor de "Depth-of-Injury" theorie getoetst zou kunnen worden. De vraag bleef of een theorie die gebaseerd is op een herstelproces in een (levend) konijnenoog dat dagen of weken kan duren, kan correleren met het schadeproces in de ICE. De uitkomst van de Draize test wordt sterk beïnvloed door verschillende oncontroleerbare factoren die niet plaatsvinden in de ICE. Deze factoren en hun consequentie voor de irritatie classificatie van stoffen worden in de discussie van de **Hoofdstukken 8 en 9** besproken.

Om de histopathologische beoordeling van de corneas te optimaliseren werd bij TNO de mogelijkheid onderzocht om meer specifieke kleuringen toe te passen bij het maken van de microscopische preparaten van de cornea (**Hoofdstuk 6**). In eerste instantie werden de microscopische preparaten van de cornea gekleurd met de standaard kleuring Haematoxylin & Eosin (H&E), maar deze werd later vervangen door de Periodic Acid-Schiff (PAS) kleuring die een beter onderscheid van de diverse lagen van de cornea gaf. In het bijzonder de diverse membranen van de cornea, het basaal membraan en Bowman's membraan (tussen epithelium and stroma) en het Descemet's membraan (tussen stroma en endothelium) waren beter te onderscheiden. Deze membranen worden geacht een belangrijke rol te spelen bij het ontstekings- en herstelproces van de cornea. Daarom was de zichtbaarheid van deze membranen belangrijk voor de vaststelling van hun integriteit bij de microscopische beoordeling door de patholoog. Andere meer specifieke kleuringen voor collageen rijke membranen zouden mogelijk de beoordeling kunnen verbeteren. Voor dit doel werden coupes van cornea's, die waren blootgesteld aan niet irriterende, irriterende of ernstig irriterende stoffen, bewerkt met de verschillende kleuringen en beoordeeld door de patholoog. Van de vijf geselecteerde kleuringen, H&E, PAS, Trichrome, AZAN (Azocarmine & aniline) and EVG (Elastic Van Gieson), bleek PAS duidelijk superieur te zijn met betrekking tot zichtbaarheid van de membranen en de kwaliteit van de morfologie van de diverse structuren van de cornea. Na ernstige corrosieve schade door natrium hydroxide

(natronloog) bleek de basaal membraan en Bowman's membraan ogenschijnlijk onbeschadigd terwijl in het onderliggende stroma wel effecten werden waargenomen. Betekende dit dat de functionaliteit van de membranen niet was aangetast of was lichtmicroscopie niet in staat dit soort schade vast te stellen? Deze observatie leidde tot de conclusie dat "Depth-of-Injury" niet de enige factor is die de ernst van de schade bepaalt. Overleg met een oogarts van het Utrecht Medisch Universiteitscentrum, gespecialiseerd in de cornea, en diverse publicaties op dit gebied wezen uit dat in de kliniek de nadruk vooral ligt op de troebeling van de cornea en stamcel overleving na (chemische) schade. Bij zeer omvangrijke schade aan de stamcellen kan de cornea niet meer herstellen door re-epithelisatie van migrerende stamcellen. In dat geval resulteert de schade in onherstelbare conjunctivale ingroei van de cornea (conjunctivalization). De stamcellen zijn oppervlakkig gelokaliseerd aan de rand van de cornea en als zodanig betrokken bij het eerste contact na blootstelling aan een stof. Daarom kan het vaststellen van de viabiliteit van de stamcellen van betekenis zijn. Pogingen om stamcellen van de kippen cornea met een p63 immuunkleuring te identificeren, bleken niet succesvol. Hoewel het vaststellen van reversibiliteit of irreversibiliteit van oogschade belangrijk is, zeker voor de producenten van huishoudelijke en verzorgingsproducten, maken de classificatie schema's van de EU en GHS geen onderscheid tussen deze twee categorieën. Ze worden beide geclassificeerd als Categorie 1: "Veroorzaakt ernstig oogletsel".

De behoefte aan geaccepteerde alternatieve testmethoden leidde tot een van de grootste validatiestudies ooit georganiseerd, namelijk de EC/HO studie in 1993-1995. Zestig stoffen werden in negen verschillende alternatieve test methoden, waaronder de ICE, onderzocht in vier verschillende laboratoria per methode. Met een maximale gemiddelde correlatie variërend van 0.34 tot 0.55 (1.0 is maximaal haalbaar) was de uitslag van deze studie erg teleurstellend. In **Hoofdstuk 7** wordt beschreven dat het gebruik van de MMAS (Modified Maximum Average Score) als enige parameter voor de mate van de *in vivo* oogirritatie, de oorzaak was van dit teleurstellende resultaat. De MMAS, die kan variëren van 0-110, is een gemiddelde van de hoogste oogweefsel scores van de individuele konijnen, vastgesteld na het testen van een stof. Een stof werd geclassificeerd als niet irriterend bij een MMAS van 0-25, irriterend bij een MMAS van 25-59 of ernstig irriterend bij een MMAS hoger dan 59. Omdat de MMAS niet gebruikt werd bij het vaststellen van de EU en GHS oogirritatie classificaties, werd onderzocht welke gevolgen deze classificaties hadden op het gebruik van de MMAS in de EC/HO studie. De MMAS grenzen van 25 en 59, behorend tot respectievelijk irriterend en ernstig irriterend, bleken niet goed overeen te komen met de EU en GHS classificaties voor irriterend en ernstig irriterend. Acht stoffen met een MMAS lager dan 59 waren ernstig irriterend en 4 stoffen met een MMAS hoger dan 59 en 3 stoffen met een MMAS lager dan 25 waren irriterend. De aanbeveling was dan ook om voortaan gebruik te maken van de EU en GHS classificaties bij nieuwe validatie studies.

In 2004 werd een nieuw initiatief gestart door het Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) in samenwerking met

het European Centre for the Validation of Alternative Methods (ECVAM) om alternatieve methoden, die het meest veelbelovend waren in de EC/HO studie, te (her)valideren voor het identificeren van ernstig irriterende stoffen. Dit waren de ICE, de Hen's Test - Chorioallantois Membrane (HET-CAM), de Isolated Rabbit Eye test (IRE) en de Bovine Corneal Opacity and (fluorescein) Penetration test (BCOP). Om de validatiestatus van deze methoden vast te stellen werd per alternatieve methode een onafhankelijk panel benoemd. In het algemeen was de gedachte dat de slechte correlatie tussen de alternatieven en de Draize test mede lag aan tekortkomingen in de praktische uitvoering van de alternatieve methoden. Aan eventuele tekortkomingen van de Draize test en de mogelijke consequenties daarvan voor de validatie van de alternatieven werd weinig aandacht besteed. Na een openbare bijeenkomst waarin experts hun alternatieve methode konden presenteren en vragen daarover beantwoorden van de panelleden, schreef Menk Prinsen een discussiestuk over dit onderwerp, getiteld: "The Draize Eye Test and *in vitro* alternatives; a left-handed marriage?" (**Hoofdstuk 8**). Hierin worden de verschillende aspecten van de Draize ogen test besproken die van invloed zijn op de ontwikkeling, validatie en acceptatie van alternatieven. Over het algemeen worden de *in vivo* data gebruikt als de "Gouden" standaard bij het vergelijken met de resultaten van de alternatieve methode, dus 100% accurate voorspelling. Hoewel men in het algemeen wel op de hoogte was van de tekortkomingen van de Draize test, werden deze nooit meegewogen in het validatie en acceptatie proces van alternatieven. Men wist dat de Draize test variabele resultaten opleverde omdat verschillende laboratoria en waarnemers bij het uitvoeren van deze test betrokken waren en de gebruikte data over een zeer lange periode waren verzameld. Maar ook andere belangrijke zaken spelen een cruciale rol bij de uitkomst van deze *in vivo* test. De blootstellingsduur aan een stof is hoegenaamd niet gedefinieerd of gestandaardiseerd. Doordat een stof in het uitgetrokken oogzakje word gedeponeerd waarna het konijn wordt teruggezet in zijn kooi, kan de blootstelling variëren van minuten (bij vloeistoffen) tot uren (24 uur bij poeders), omdat resten van stoffen pas 24 uur na blootstelling verwijderd mogen worden. Deze ongedefinieerde blootstelling is tegenstrijdig met de basis principes van goed toxiciteitsonderzoek; alles valt of staat bij een goed gedefinieerde en gecontroleerde blootstelling. Daarbij, dit type blootstelling (grote hoeveelheid stof in een oogzakje) is nauwelijks vergelijkbaar met een humane blootstelling. In plaats van dit ter discussie te stellen, concludeerde ICCVAM dat de ICE in dit opzicht tekort schoot en kwam met de aanbeveling om de ICE te optimaliseren met betrekking tot de blootstellingprocedure voor vaste stoffen. Een ander onderbelichte factor vormde de verzorging van het oog na blootstelling. De opsluiting van bijvoorbeeld vaste stoffen in het oogzakje kan leiden tot een volledige afsluiting van het oog door verkleving van de oogleden vanwege de productie van ontstekingseiwitten, die een afsluitende korst vormen.
De ervaring was dat, als het oog en de oogleden niet regelmatig werden verzorgd (schoonmaken/spoelen met een fysiologische zoutoplossing), een irriterende stof gemakkelijk kon veranderen in een ernstig irriterende stof. Een dergelijke, in wezen, minimale verzorging van het blootgestelde oog en de oogleden vond in het algemeen

nooit plaats en werd ook niet genoemd in de richtlijnen. Deze omstandigheden zijn ondenkbaar in het geval van humane blootstelling aan irriterende stoffen. Een ander fenomeen dat kon plaatsvinden in de Draize test was het optreden van een secundaire ooginfectie. Door de irriterende effecten kan een oog verzwakt raken en gevoelig worden voor microbiële infecties. Daardoor kunnen in eerste instantie matige oogeffecten veranderen in ernstige aanhoudende effecten gedurende de observatieperiode van 21 dagen. Deze gevallen van misinterpretatie van effecten, waarvoor al in de eerste OECD richtlijn voor oogirritatie werd gewaarschuwd, zijn nooit in overweging genomen door instanties zoals ICCVAM en ECVAM, ondanks dat er voldoende aanwijzingen waren voor zo'n infectie in de gebruikte *in vivo* data set.
De conclusie van het discussiestuk was dat na 18 jaar van validatie zonder daadwerkelijk succes, het verder ontwikkelen en valideren van alternatieven zinloos zou zijn als men gebruik bleef maken van de bestaande *in vivo* data of nieuw te genereren *in vivo* data met de bestaande richtlijn. Een meer-sporen beleid werd aanbevolen, waarbij de bestaande OECD richtlijn voor oogirritatie de blootstelling aan vloeibare en vaste stoffen zou standaardiseren en het gebruik van alternatieven voor ernstige oogirritatie zou opnemen (dit werd overigens al toegestaan binnen de EU). Parallel daaraan zouden de *in vitro* methoden meer op mechanistische grondslag ontwikkeld en geoptimaliseerd moeten worden.

In **Hoofdstuk 9** worden de resultaten samengevat, toegelicht en bediscussieerd. Daarbij wordt nog dieper ingegaan op de onmogelijkheid om van de alternatieve methoden een 100% correlatie te verlangen met de "onbetrouwbare" *in vivo* test.
In 1971 werd door Amerikaanse onderzoekers vergelijkend *in vivo* oogirritatieonderzoek uitgevoerd door dezelfde stoffen te laten testen door een twintigtal laboratoria. Op basis van die resultaten werd toen al geconcludeerd dat de Draize test niet geschikt was om door regulerende instanties als "gouden standaard test" beschouwd te worden. Desondanks is de Draize test wereldwijd de test die gebruikt moet worden om de mogelijke oog irriterende werking van stoffen te onderzoeken. In 2009 kwam er eindelijk een doorbraak door de acceptatie van de ICE en BCOP door de OECD als richtlijn voor het screenen van stoffen op ernstige oogirritatie. Ondanks het feit dat er een aanzienlijk percentage vals-negatieven was (niet ernstig irriterend in de *in vitro* test maar ernstig irriterend in de *in vivo* test) werd dit niet als een belemmering gezien omdat de niet ernstig irriterende stoffen *in vitro* altijd nog onderzocht moesten worden in de *in vivo* test. Veel van de zogenaamde vals-negatieven in de ICE test laten *in vivo* resultaten zien die verklaard kunnen worden met de beschreven tekortkomingen van de Draize test. In 2011 volgde een nieuw initiatief om ook het screenen van niet-irriterende stoffen door de BCOP op te nemen in de OECD richtlijn van 2009. De ICE werd niet geschikt geacht, voornamelijk omdat een tweetal stoffen die ernstig irriterend waren in de *in vivo* test niet irriterend waren in de ICE. Naar aanleiding daarvan werd door Menk Prinsen de gebruikte dataset voorzien van inhoudelijk commentaar met betrekking tot tekortkomingen of bijzonderheden in de *in vivo* data en ingestuurd met verzoek tot heroverweging van de ICE. Daaropvolgend werd bij de OECD een expert meeting

georganiseerd waar Menk Prinsen een presentatie gaf die toegespitst was op de problemen die spelen bij de *in vivo* test en de invloed daarvan op de resultaten. Naar aanleiding daarvan werden deze punten opgenomen in een OECD "Streamlined Summary Document" van de ICE en de resultaten na re-evaluatie geaccepteerd (82% accuratesse, 1% vals-negatieven en 33% vals-positieven). De BCOP werd geaccepteerd met 69% accuratesse, 0% vals-negatieven en 69% vals-positieven. In september 2013 werd het screenen van niet-irriterende stoffen door de ICE officieel toegevoegd aan de OECD richtlijn van de ICE. Dit was een groot succes voor de methode en voor de 3V's in het algemeen, maar betekende nog wel dat stoffen die irriterend waren in de ICE (of BCOP) nog steeds *in vivo* getest moeten worden. In de praktijk blijkt dat de branche die het meest te maken heeft met irriterende producten, namelijk die van de huishoudelijke- en industriële schoonmaakmiddelen, om te voldoen aan de regelgeving, al een dierproefvrije strategie toepast bij oogirritatieonderzoek. Het delen van informatie over deze dierproefvrije strategie met andere branches en industrieën is één van de aanbevelingen in dit proefschrift.

De eindconclusie van dit proefschrift is dat alternatieven voor de Draize oogirritatie test in konijnen bij voorkeur gebruik zouden moeten maken van geïsoleerde ogen, oogweefsel (geïsoleerde cornea) of oogweefsel equivalenten (gereconstrueerd cornea epithelium) om zowel kwalitatief als kwantitatief parameters te kunnen meten die gelijkwaardig of identiek zijn aan parameters voor humane oogschade. Modellen die gebruik maken van geïsoleerde "intacte" ogen met waarnemingen door middel van een oogmicroscoop en met de mogelijkheid om histopathologie toe te passen, zoals de ICE en IRE, worden daarvoor het meest geschikt geacht.

Appendices

Dankwoord
Curriculum vitae
List of publications
Overview of training activities

Dankwoord

Dertig jaar met een alternatieve test voor oogirritatie bezig zijn en meer dan 40 jaar bij TNO werken, dan wordt het lastig om iedereen te bedanken die een rol van betekenis heeft gespeeld in mijn loopbaan bij TNO. Toch ga ik een poging wagen en direct bij aanvang mijn verontschuldigingen aanbieden aan alle behulpzame mensen die ik vergeten ben te noemen in dit dankwoord.

Ik begin met de twee personen die het belangrijkst zijn geweest voor mijn ontwikkeling bij TNO, namelijk Herman Koëter en Maarten Bosland. Met Maarten heb ik als onderzoeksanalist samen gewerkt aan zijn onderzoeksproject naar prostaatkanker voor het KWF van 1980-1984. Zowel op werkgebied als privé konden we het goed met elkaar vinden en ik denk met ontzettend veel plezier terug aan die tijd, waarin ik zoveel geleerd heb op het gebied van onderzoek doen. Maarten gaf mij alle vrijheid en vertrouwen om mee te denken en voorstellen te doen voor het praktische gedeelte van het onderzoek. Het heeft mede de basis gelegd voor mijn ontwikkeling tot toxicoloog. Datzelfde geldt voor mijn samenwerking met Herman.

Herman, jij was mijn voorbeeld voor wat je kunt bereiken, wanneer je bereid bent om naast je werk alle energie in je ontwikkeling te steken. Bij jou begon ik als biotechnicus bij het teratologisch onderzoek dat je bij TNO aan het opzetten was. Na het behalen van mijn zoölogisch analisten diploma kreeg ik de kans om studie director te worden bij de sectie Reproductie en Dermale Toxicologie waarvan jij toen hoofd was geworden.
Begin jaren tachtig maakte Herman het mogelijk dat er een alternatief voor de oogirritatie test bij konijnen op TNO werd geïntroduceerd. Vanaf het moment dat er besloten was om geïsoleerde ogen te gebruiken als alternatief, kreeg ik van hem alle vrijheid en vertrouwen om de methode op te zetten en te toetsen. Zijn vertrek naar de OECD begin jaren 90 betreurde ik vanwege het verlies van een mentor, maar het gaf mij ook de gelegenheid om de vleugels uit te slaan en de volledige verantwoordelijkheid voor de methode op mij te nemen.

Bij het eigen maken van de *in vivo* huid- en oogirritatie studies als studie director werd ik ingewerkt door Lammert van Beek, destijds verantwoordelijk voor de dermale toxiciteitsstudies. Hij heeft mij alle “ins” en “outs” van de *in vivo* studies bijgebracht, met grote aandacht voor het beoordelen van de blootgestelde huiden en ogen van de konijnen, iets waarvoor ik hem altijd nog zeer erkentelijk ben. Sinds die tijd heb ik samengewerkt met verschillende dierverzorgers en biotechnici, waarvan Theo Woertman en Herman van Vulpen die van het eerste uur waren. Zij en alle andere dierverzorgers en biotechnici, die mij sindsdien hebben geassisteerd bij de uitvoering en beoordeling van huid- en oogirritatietesten, wil ik hierbij bedanken. De kennis die ik daarbij heb opgedaan was essentieel bij het ontwikkelen van de ICE test en voor mijn ontwikkeling als deskundige op het gebied van oogirritatie.

Reier van Barneveld van de (opgeheven) Technische Dienst van TNO heeft een grote bijdrage geleverd door het vervaardigen de eerste twee versies van de ooghouders en het superfusie apparaat voor de ogen. Nu 30 jaar later gebruik ik deze ooghouders nog steeds.
De slachthuizen van de Bor in Nijkerkerveen en van de Miert in Breukelen ben ik zeer erkentelijk voor het beschikbaar stellen van de kippenkoppen al die jaren.

Lidy (helaas overleden in juli dit jaar), Darryl, Sebastiaan, Tim, Greetje, Lisanne en Hannie van de sectie Histologie van TNO, jullie bedank ik voor het verwerken van de ogen tot coupes. Tim, speciaal bedankt voor het transport van de kippenkoppen vanaf de slachthuizen. Darryl mijn dank voor de snelle potjes service en het excel(lente) verwerken van de data. Voor de microscopische beoordeling van de coupes van de cornea was de bijdrage van Joost Bruijntjes en Marcel Wijnands onmisbaar. Dit onderdeel is een vaste waarde geworden in de eindbeoordeling van de stoffen. Marcel, jou wil ik speciaal bedanken voor het werk aan de atlas met de karakterisering van histopathologische afwijkingen van de cornea in de ICE test.

Naast de ICE test heb ik ook meegedaan aan de validatie van een ander alternatief voor oogirritatie, de BCOP test. Wilfred Maas wil ik bedanken voor al het praktische werk dat hij daaraan heeft verricht. Als toenmalig (armlastig?) student was de leenauto, die je voor het ophalen van de runderogen in Den Bosch kreeg, natuurlijk een prettige bijeenkomst. De ervaring die we opdeden met deze test was ook waardevol voor het op de juiste waarde schatten van de ICE.

Ik ben nog lang niet klaar met bedanken. Ik wil graag Rob Roggeband noemen, die de ICE introduceerde bij de Procter and Gamble Company waar hij werkzaam is. Zonder zijn vertrouwen in de mogelijkheden van de ICE denk ik niet dat deze methode zover was gekomen als nu het geval is. Verder zijn er de collega's van P&G die zich bezighouden met oogirritatie, met name Pauline McNamee en Katrin Schutte, met wie ik zeer prettig heb samengewerkt.

De organisaties ECVAM en ICCVAM ben ik erkentelijk voor de manier waarop zij zich hebben gekweten van de moeilijke taak om alternatieven door het mijnenveld, dat validatie is gebleken, te loodsen, zeker op het gebied van de Draize oogirritatie test. Om heelhuids daar doorheen te komen is een cursus politiek en diplomatie voor onderzoekers zeer aan te raden.

I warmly thank the organization AISE, her members and especially Elodie Cazalle for their cooperation in the further development and application of histopathology in the ICE test in the domain of detergents.

Chantra Eskens, thank you so much for your work and effort as a consultant for the OECD in preparing and presenting the ICE dossier for the acceptance and adoption as an OECD guideline for the identification of non-irritants. Your diplomacy set an example for me.

Er zijn nog een aantal personen die ik graag wil noemen:
Betty Hakkert, nationaal coördinator voor Nederland bij de OECD vanwege haar raad en ondersteuning en haar inspanningen om de ICE test geaccepteerd te krijgen voor de identificering van niet-irriterende stoffen door de OECD.

Peter Davis voor het beoordelen van de hoofdstukken “Introduction” en “Discussion” met betrekking tot de Engelse grammatica.

Mijn promotoren Ruud en Coenraad en mijn copromotor Cyrille voor hun inbreng en discussies bij de totstandkoming van dit proefschrift. Ruud, jou ben ik zeer erkentelijk voor de vrijheid die jij mij gaf om het onderzoek naar eigen inzicht in te vullen tijdens de periode dat je hoofd van de afdeling TAP was. Ook motiveerde jij mij steeds om te gaan (en door te gaan met) promoveren op dit onderwerp.

Ik wil Alfred van Rossum bedanken voor zijn inspanningen en geduld en bij het grafisch ontwerpen van dit boekje. Alfred, ik hoop niet dat ik je lastigste klant ben geweest in al die jaren als grafisch ontwerper.

Micha Prinsen en Iwan, ontzettend bedankt voor jullie mooie ontwerp voor de omslag van het boekje. Veel succes verder met jullie kunstenaarschap en bierbrouwen (Rotterdamse Bocht) in Roffa.

Lieve Cindy, nu ben jij dan eindelijk aan de beurt. Bedankt dat je al die “TNO” jaren met mij hebt doorstaan. Bij het schrijven van dit dankwoord realiseerde ik mij opeens dat onze relatie en dit onderzoek eigenlijk precies gelijk oplopen. Beide, oneerbiedig genoemd, projecten kenden hun “ups” en “downs”, maar zijn zeer succesvol gebleken, met natuurlijk als mooiste resultaat onze drie kinderen, Joeri, Sosja en Micha. Aan verdere vergelijkingen zal ik mij niet meer wagen. Het zal je, afkomstig uit een vegetariër nest, niet altijd gemakkelijk gevallen zijn om met iemand samen te leven die al vanaf zijn 18de met proefdieren werkt en zeer rationeel denkt. Dat ik altijd zeer van de controle en vooruitdenken ben valt niet altijd mee, maar je hebt mij altijd alle ruimte gegeven. Wat ik heb bereikt op mijn werkgebied is voor een groot deel aan jouw ondersteuning te danken. Jouw zorg voor het thuisfront tijdens de kinderjaren maakte het voor mij mogelijk om deze carrière bij TNO op te bouwen. Fantastisch dat je, nadat Micha naar de middelbare school ging, zelf nog een succesvolle zaak (Stoel 67) hebt opgebouwd. Joeri, Sosja en Micha, jullie zijn als laatste aan de beurt. Vader van jullie te zijn, is het beste wat mij is overkomen en heeft mij in al die jaren (natuurlijk met de nodige kopzorgen) veel over mijzelf geleerd. Jullie zien opgroeien tot de volwassenen die jullie nu zijn, was en is een heel bijzondere ervaring. Enfin, jullie moeten nu wel verder met een Doctor.

Menk K. Prinsen
Utrecht, October 2014

&

Curriculum Vitae

Menk Prinsen werd geboren op 30 juli 1954 te Amersfoort. Na zijn middelbare schoolopleiding ging hij naar de Amersfoortse Laboratoriumschool en koos na het eerste algemene jaar voor een 3-jarige (avond)opleiding tot dierverzorger/biotechnisch laborant aan de Dr. Ir. W.L. Ghijsen Instituut te Utrecht. Tijdens die 3 jaren was hij eerst werkzaam als leerling dierverzorger op het Centraal Proefbedrijf van TNO te Austerlitz en vervolgens als dierverzorger en biotechnisch laborant bij het Centraal Instituut voor Voedingsonderzoek (CIVO) van TNO te Zeist. Na het behalen van het diploma biotechnisch laborant ging hij terug naar de Amersfoortse Laboratoriumschool voor het voorbereidende jaar HBO. Daarna keerde hij terug bij het CIVO en voltooide de (avond)opleiding Zoölogisch analist. Gedurende die periode was hij betrokken bij o.a. teratologisch onderzoek en later bij het KWF onderzoek naar de rol van voeding bij het ontstaan van prostaatkanker, waarvoor een specifiek model in ratten werd ontwikkeld.

In 1981 volgde een interne opleiding tot studie director van huid- en oogirritatie onderzoek dat later werd uitgebreid met acuut oraal en acuut en sub-chronisch dermaal toxiciteitsonderzoek, dermale (foto)sensibilisatie, fototoxiciteit en testen op uterotrophe activiteit van stoffen. Met de ervaring die hij daarbij opdeed was hij betrokken als deskundige bij het voorbereiden en opstellen van diverse OECD richtlijnen. Vanaf 1990 was hij studie director van preklinische “safety pharmacology” studies en meer specifiek van “vaccin safety” studies. Na het instellen van de erkenning toxicoloog via de toegepaste route door de Nederlandse Vereniging van Toxicologen volgde hij de postdoc opleidingen Risico-evaluatie, Principles of Toxicological Pathology en Medische en Forensische toxicologie en verkreeg de erkenning tot toxicoloog.

Vanaf 1982 tot op heden houdt hij zich, naast zijn taak als studieleider van bovengenoemde toxiciteitstesten, bezig met het ontwikkelen en implementeren van de geïsoleerde ogen test als volwaardig dierproefvrij alternatief voor de Draize oogirritatietest met konijnen. In 2010 ontving hij voor dit werk van de Stichting Bouwstenen voor Dierenbescherming de “Hugo van Poelgeest” prijs.

List of publications

Bosland, M.C., M.K. Prinsen and R. Kroes (1983). Adenocarcinomas of the prostate induced by N-nitrosomethyl urea in rats pretreated with Cyproterone acetate and testosterone. *Cancer Letters*, 18: 69-78.

Bosland, M.C., Prinsen, M.K., Dirksen, T.J.M. and Spit, B.J. (1990). Characterization of Adenocarcinomas of the Dorsolteral Prostrate Induced in Wistar Rats by N-Methyl-N-nitrosourea, 7,12-Dimethylbenz(a)anthracene, and 3,2'-Dimethyl-4-amino-biphenyl, following Sequential Treatment with Cyproteropne Acetate and Testosterone Propionate. *Cancer Research*, 50: 700-709.

Bosland, M.C. and Prinsen, M.K. (1990). Induction of Dorsolateral Protstate Adeno-carcinomas and Other Accessory Sex Gland Lesions in Male Wistar Rats by a Single Administration of N-Methyl-N-nitrosourea, 7,12-Dimethylbenz(a)anthracene, and 3,2'-Dimethyl-4-animopiphenyl after Sequential Treatment with Cyproterone Acetate and Testosterone Propionate. *Cancer Research*, 50: 691-699.

Bosland, M.C., Menk K. Prinsen, Abraham Rivenson and John H. Weisburger (1992). Induction of skin and thyroid tumors in male rats by N-methyl-N-nitrosourea after sequential treatment with cyproterone acetate and testosterone propionate: effects of castration, rat strain and time of carcinogen injection. *Carcinogenesis*, Volume 13, no. 4, 669-674.

Bunnik, G.S.J., M.C. Bosland, M.K. Prinsen and E. Hidajat (1982). Dietary influences on plasma levels of testosterone, 5>-dihydrotestosterone and prolactin in male rats. *J. Steroid Biochem.*, 17: LXXXIII, abstr. 249.

Fentem, J.H., M.K. Prinsen, H. Spielman, E. Walum and P.A. Botham (1990). Validation - lessons learned from practical experience. *Toxicology In Vitro*, Volume 9, no. 6, pp. 857-862.

Gautheron, P. and 17 co-authors (including M.K. Prinsen) (1994). Interlaboratory assessment of the Bovine Corneal Opacity and Permeability (BCOP) assay. *Toxicology In Vitro*, Volume 8, no. 3, pp. 381-392.

Koëter, H.B.W.M., Prinsen, M.K. (1985). Comparison of in vivo and in vitro eye irritation test systems: A study with 34 substances. *Alternative Methods in Toxicology* 3, Chapter A9. Mary Ann Liebert, Inc., publishers, New York.

Koëter, H.B.W.M., Prinsen, M.K. (1987). Validation of an in vitro eye irritation study; A first step. *Alternative Methods in Toxicology* 5, Chapter E5. Mary Ann Liebert, Inc., publishers, New York.

&

Prinsen, M.K., Koëter, H.B.W.M. (1993). Justification of the Enucleated Eye Test with eyes of slaughterhouse animals as an alternative to the Draize eye irritation test with rabbits. *Food and Chemical Toxicology* 31(1), 69-76.

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 and Chemical Toxicology* 34(4), 291-296.

Til, H.P., H.E. Falke, M.K. Prinsen and M.I. Willems (1997). Acute and subacute toxicity of Tyramine, Spermidine, Spermine, Putrescine and Cadaverine in rats. *Food and Chemical Toxicology* 35: 337-348.

Hempenius, R.A., J.H.M. van Delft, M. Prinsen and B.A.R. Lina (1997). Preliminary Safety Assessment of an Arochidonic Acid-enriched Oil derived from Mortierella alpina: Summary of Toxicological Data. *Food and Chemical Toxicology* 35: 573-581.

Prinsen, M.K., T. Romijn and N.J. Snoeij (1997). Skin sensitization testing the relevance of rechallenge and pretreatment with sodium lauryl sulfate in the guinea pig maximization test. *Food and Chemical Toxicology*, Volume 35, pp. 923-926.

Prinsen, M.K. and N.J. Snoeij (1999). Skin sensitization testing the relevance of rechallenge and pretreatment with sodium lauryl sulfate in the guinea pig maximization test. Answer to a Letter to the Editor. *Food and Chemical Toxicology*, Volume 37, pp. 265-269.

Prinsen, M.K. (1999). An Evaluation of the OECD Proposal for the Harmonised Classification of Eye Irritants and Corrosives. Report of ECVAM Workshop 34, Eye Irritation Testing: The Way Forward, Appendix 1. *ATLA* 27, 72-77.

Prinsen, M.K. and N. Gouko (2001). Determination of the Oestrogenic (Uterotrophic) Activity of Extracts of "General Purpose Polystyrene (GPPS)" using Immature Female Rats. *Journal of Applied Toxiciolgy* 21, 235-239.

Emmen, H.H., H. Muijser, J.H.E. Arts, M.K. Prinsen (2003). Human volunteer study with PGME; eye irritation during vapour exposure. *Toxicology Letters* 140-141: 249-259.

Prinsen, M.K. (2006). The Draize Eye Test and in vitro alternatives; a left-handed marriage? *Toxicology In Vitro* 20, 78-81.

Schutte, K., M.K. Prinsen, P.M. McNamee, R. Roggeband (2009). The isolated chicken eye test as a suitable in vitro method for determining the eye irritation potential of household cleaning products. *Regulatory Toxicology and Pharmacology* 54, 272–281.

Prinsen, M.K., Dr. M.E.I. Schipper and Dr. M. Wijnands (2011).The role of histopathology in the isolated chicken eye test and comparison of different stainings of the cornea. *Toxicology In Vitro* 25, 1475-1479.

Cazelle, E., Eskes, C., Hermann, M., Jones, P., McNamee, P., Prinsen, M., Taylor, H., Wijnands, M.V.W. (2014). Suitability of Histopathology as an Additional Endpoint to the Isolated Chicken Eye Test for Classification of non-extreme pH Detergent and Cleaning Products. *Toxicology In Vitro* 28, 657-666.

Overview of training activities

a. Discipline specific activities	Year
ICCVAM Ocular Toxicity Working Group meeting (invited speaker), NIEH, Bethesda, USA	2004
Expert meeting on alternatives for eye irritation (presentation ICE), NIEH, Bethesda, USA	2005
CAAT World Congress, Berlin, Germany	2005
DIA 1-day Workshop on non-clinical testing of Vaccines, Amsterdam	2007
DIA 1-day Workshop on Photosafety Evaluation of Drugs, Amsterdam	2007
Minipig Research Forum, Copenhagen, Denmark	2007
Training course (2 weeks) ICE of staff member University of Durban, Zeist	2008
1st International Conference on Dermatotoxicology, Vaalsbroek	2008
Minipig Research Forum, Copenhagen, Denmark	2008
CBER/NIAID Adjuvant Workshop, Bethesda MD, USA	2008
OECD Expert Consultation Meeting on Ocular Irritation/Corrosion (presentation ICE), Washington, USA	2008
15th International Congress on Photobiology, Düsseldorf, Germany	2009
CAAT Poster presentation ICE staining of corneas, World Congress, Rome, Italy	2009
Presentation ICE award ceremony: "Hugo van Poelgeestprijs", Poelgeest	2010
ICE 5-day Training course of staff University of Durban, South-Africa	2011
Key lecture Alternatives, SAALAS Congress, Johannesburg, South-Africa	2011
VWA workshop. ICE presentation, Den Hague	2011
NCV; ICE presentation, Houten	2011
NVT sectie Risicobeoordeling, ICE demonstration and presentation, Zeist	2011
Annual lecture ICE, Hogeschool Utrecht	2007-2011
Presentation/discussion OECD guideline 438, Paris, France	2012
BELTOX 1-day workshop vaccines, Wavre, Belgium	2014
b. General courses/activities	
Post-doc education Toxicology, Risk Evaluation and Risk Assessment, Wageningen	1999
Post-doc education Toxicology, Medical and Forensic Toxicology, Utrecht	2000
Modular training programme, Principles of Toxicological Pathology, University of Surrey, UK	2001
Modular training programme, Haematology and clinical biochemistry, University of Surrey, UK	2004
GLP Training Seminar Refresher Study Director Training, Zeist	2006-2012
c. Optionals	
Preparation PhD research proposal	
Discussion meetings pathologist ICE, Zeist	2008-2014
Discussion meetings ophthalmologist, Zeist	2009-2014

Approved by the graduate school VLAG

List of abbreviations

AZAN	Azocarmine & Aniline
BCOP	Bovine Corneal Opacity and (fluorescein) Penetration test
CEET	Chicken Enucleated Eye test
CRO	Contract Research Organization
EC	European Community
ECVAM	European Centre for the Validation of Alternative Methods
EU	European Union
EVG	Elastic Van Gieson
FDA	Food and Drug Administration
H&E	Haematoxylin & Eosin
HET-CAM	Hen's Test - Chorioallantois Membrane
HO	British Home Office
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
ICE	Isolated Chicken Eye test
IRAG	Interagency Regulatory Alternatives Group of the United States
IRE	Isolated Rabbit Eye test
MMAS	Modified Maximum Average Score
NEI	National Eye Institute, USA
OECD	Organisation for Economic Co-operation and Development
PAS	Periodic Acid-Schiff
PM	Prediction Model
REET	Rabbit Enucleated Eye test
TG	Test Guideline
TNO-CIVO	Toegepast Natuurwetenschappelijk Onderzoek - Centraal Instituut voor Voedingsonderzoek
UN-GHS	United Nations - Globally Harmonized System

Financial support from Wageningen University for printing this thesis is gratefully acknowledged.

Printing: GVO drukkers & vormgevers B.V. | Ponsen & Looijen
Layout: Alfred van Rossum, Utrecht, the Netherlands
Cover design: Iwan Smit (iwansmitart@gmail.com) and Micha Prinsen (michaprinsen@gmail.com), Rotterdam, the Netherlands