Background Review Document of an *In Vitro* Approach for EPA Toxicity Labeling of Anti-Microbial Cleaning Products

Prepared for:

Alternatives Testing Steering Committee
Johnson Diversey
S.C. Johnson & Son, Inc.
The Procter & Gamble Company
The Accord Group

Prepared by:

Institute for In Vitro Sciences, Inc.
30 West Watkins Mill Road, Suite 100
Gaithersburg, MD 20878
USA

Authors:

Rodger D. Curren, Ph.D.
Jennifer R. Nash, M.S.
Angela Sizemore, B.S.
John Harbell, Ph.D.

Table of Contents

List of Tables vi

List of Figures xvi

Annexes xxi

List of Abbreviations xxii

Acknowledgements xxiii

Preface xxiv

Executive Summary xxivi

1 Introduction and Rationale for the Proposed Test Method 1

1.1 Introduction 1
1.1.1 Description of framework for development of program 1
1.1.2 Summary of Project History 2
1.1.3 Confidential information 7
1.2 Regulatory rationale and applicability 7
1.2.1 Current regulatory testing requirements for which the proposed test method is applicable 7
1.2.2 Intended regulatory use (*i.e.,* replacement) of the proposed method 10
1.2.3 Similarities between data obtained using this method and the current *in vivo* data 10
1.2.4 Fit of method into the overall strategy of toxicity or safety assessment 10
1.3 Scientific basis for the proposed test method 10
1.3.1 Purpose and mechanistic basis of the proposed test methods 10
1.3.1.1 Cytosensor Microphysiometer (CM) Assay 10
1.3.1.1.1 Intended uses / purpose of the CM 11
1.3.1.1.2 Regulatory rationale and applicability of the CM 11
1.3.1.1.3 Scientific basis for the CM test 11
1.3.1.2 EpiOcular 16
1.3.1.2.1 Intended uses / purpose of the EpiOcular assay 17

1.3.1.2.2 Regulatory rationale and applicability of the EpiOcular test method 17
1.3.1.2.3 Scientific basis for the EpiOcular test method 18
1.3.1.3 BCOP 20
1.3.1.3.1 Intended uses / purpose of the BCOP assay 20
1.3.1.3.2 Regulatory rationale and applicability of the BCOP test method 21
1.3.1.3.3 Scientific basis for the BCOP method 21

2 Test Method Components 24

2.1 Overview of the proposed testing approach 24
2.2 Detailed description and rationale for each assay 28
2.2.1 Overview of how the CM test method is conducted 28
2.2.1.1 Development of Conversion Algorithm between SM and CM 31
2.2.2 Overview of how the EpiOcular test method is conducted 33
2.2.2.1 Preparation of the EpiOcular tissue (Description provided by the manufacturer, MatTek Corporation, Ashland, MA) 33
2.2.2.2 Test methodology 34
2.2.3 Overview of how the BCOP test method is conducted 37
2.3 Use of histology in conjunction with the BCOP assay 42

3 Substances Used For Validation of the Proposed Testing Approach 43

3.1 Rationale for the products selected, including rationale for solicitation of additional test materials to fill in gaps 43
3.2 Rationale for dividing substances into “buckets” 43
3.3 Rationale for number of substances included in the study 44
3.4 Chemicals or products evaluated 45
3.5 Coding procedures 45

4 *In vivo* Reference data used for the assessment of accuracy 46

4.1 Protocols used to generate the in vivo data 46
4.1.1 Draize rabbit eye irritation protocol 46
4.1.2 LVET rabbit eye irritation protocol 48
4.1.3 Comparison of Draize and LVET 48
4.2 Original reference data 51
4.3 Description of EPA toxicity categories 52
4.4 Description of GHS toxicity categories 53
4.5 Transformation of original data to toxicity categories 55

4.6 Quality of in vivo data 55
4.7 Human toxicity information on cleaning products 55
4.7.1 Clinical Studies by Beckley *et al.* (1965) on a light duty liquid detergent 56
4.7.2 Clinical Studies by Beckley *et al.* (1969) on a soap suspension and a liquid household cleaner 57
4.7.3 Clinical Studies by Ghassemi, *et al.* (1997) on a liquid household cleaner 58
4.7.4 Clinical studies of liquid detergent products by Roggeband, *et al.* (2000) 59
4.8 Accuracy and reliability of the LVET and Draize tests 61
4.8.1 Analysis of six rabbit tests in combinations of three 62
4.8.2 Historic references on reliability of the Draize test 65

5 Test method data and results 66

5.1 Description of the test method protocols used to generate data 66
5.1.1 Cytosensor method 66
5.1.2 EpiOcular method 67
5.1.3 BCOP method 67
5.2 Availability of copies of original data used to evaluate the predictive capacity and reliability of the three test methods 68
5.2.1 Cytosensor data 68
5.2.2 EpiOcular data 68
5.2.3 BCOP data 68
5.3 Summary of results and prediction models used to evaluate the data 69
5.3.1 Cytosensor test method 69
5.3.1.1 Company Cytosensor data submissions paired with data from the LVET assay 69
5.3.1.2 CTFA Phase III study (Gettings, Lordo et al. 1996) 75
5.3.1.3 COLIPA Validation study for eye irritation 78
5.3.2 EpiOcular 80
5.3.3 BCOP 82
5.3.3.1 Data from participating companies 82
5.4 Use of coded chemicals and compliance with GLP Guidelines 84
5.4.1 Company-submitted anti-microbial cleaning product *in vitro* data 84
5.4.2 Data obtained from secondary sources 84

6 Test Method Predictive Capacity 85

6.1 Cytosensor predictive capacity 85
6.1.1 Using the LVET assay to define a prediction model for the CM 85
6.1.1.1 Secondary analysis of acidic and alkaline materials 91
6.1.2 Using the Draize assay to define a prediction model for the CM 95
6.1.2.1 CTFA Phase III Evaluation 95

6.1.2.2 COLIPA Evaluation 99
6.1.3 Cytosensor studies without animal data 103
6.1.4 Conclusion for the Cytosensor assay 104
6.2 EpiOcular predictive capacity 106
6.2.1 Company submissions 106
6.2.2 Conclusion for EpiOcular studies 121
6.3 BCOP predictive capacity 123
6.3.1 Overview 123
6.3.2 Analysis using only BCOP *in vitro* scores (no histopathology) 123
6.3.2.1 Original company data submissions 123
6.3.2.2 Further analysis 126
6.3.2.2.1 Additional materials tested and analyzed by EPA category 126
6.3.2.2.2 Additional materials tested and analyzed by GHS toxicity category 128
6.3.2.2.3 Analysis of anti-microbial cleaning formulations with high solvent concentrations 130
6.3.3 Histopathology Analysis 135
6.3.3.1 Analysis of the predictive capacity of BCOP including histological evaluation for EPA hazard classifications 138
6.3.3.2 Analysis by GHS category for BCOP including histological evaluation 141
6.3.3.3 Conclusions from analysis of the BCOP predictive capacity 143
6.4 Strategic approach 146

7 Test Method Reliability 147

7.1 Cytosensor 149
7.1.1 Cytosensor intralaboratory repeatability I 149
7.1.2 Cytosensor intralaboratory reproducibility II 154
7.1.3 Cytosensor interlaboratory reproducibility 157
7.2 EpiOcular model 164
7.2.1 EpiOcular intralaboratory repeatability for antimicrobial cleaning products submitted by participating companies (within run and between experiments) 164
7.2.2 EpiOcular intralaboratory reproducibility (between experiments) 166
7.2.3 EpiOcular interlaboratory reproducibility 167
7.3 BCOP assay 171
7.3.1 BCOP intralaboratory repeatability 171
7.3.1.1 BCOP within-run reproducibility for antimicrobial cleaning products data 171
7.3.1.2 BCOP within-run reproducibility for a wide range of materials 190
7.3.2 BCOP intralaboratory reproducibility 191
7.3.2.1 BCOP intralaboratory reproducibility for antimicrobial cleaning products data 191

7.3.2.2 BCOP intralaboratory reproducibility for a wide range of materials 192
7.3.3 BCOP interlaboratory reproducibility 192

8 Test Method Data Quality 198

8.1 Adherence to National and International GLP Guidelines 198
8.2 Data Quality Audits 198
8.3 Impact of Deviation from GLP Guidelines 198
8.4 Availability of Laboratory Notebooks or Other Records 198

9 Other Scientific Reports and Reviews 199

10 Animal Welfare Considerations 200

10.1 How the proposed non-animal testing strategy will refine, reduce or replace animal use for the purpose of toxicity labeling of anti-microbial cleaning products 200

11 Practical Considerations 201

11.1 Use by industry 201
11.2 Ease of transferability 203
11.2.1 Facilities and major fixed equipment for the Cytosensor test method 203
11.2.2 Facilities and major fixed equipment for the EpiOcular test method 204
11.2.3 Facilities and major fixed equipment for the BCOP test method 204
11.3 Training required 205
11.3.1 Required level of training and expertise needed to conduct the Cytosensor assay 205
11.3.2 Required level of training and expertise needed to conduct the EpiOcular assay 205
11.3.3 Required level of training and expertise needed to conduct the BCOP assay 205
11.4 Cost Considerations 206
11.5 Time Considerations 207
11.5.1 Timing for Cytosensor test method 207
11.5.2 Timing for EpiOcular test method 207
11.5.3 Timing for BCOP test method 208

12 References 209

List of Tables

Table 1-1 *In vivo* Ocular Irritancy Classification Systems 8

Table 1-2 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses..... 15

Table 1-3 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses..... 19

Table 1-4 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses..... 23

Table 2-1 BCOP *in vitro* score and EPA category designation..... 25

Table 2-2 Silicon Microphysiometer data for 11 surfactant-containing materials from P&G 32

Table 2-3 Cytosensor Microphysiometer data for 11 surfactant-containing materials from P&G 32

Table 3-1 Descriptive subcategory of products tested in the individual assays. Final graphs may contain fewer materials as final applicability domains were determined. ... 45

Table 4-1 Scale of weighted scores for grading the severity of ocular lesions (Draize, Woodard et al. 1944). 47

Table 4-2 Mean time to clear after direct instillation of household cleaning products to both rabbits and humans. Compiled from Freeberg *et al.* 1986. 50

Table 4-3 Rabbit and human eye responses after exposure to either 100 µL (Draize protocol) or 10 µL (LVET protocol). All scoring done by the traditional Draize scoring scale. Compiled from Freeberg *et al.* (1986) 50

Table 4-4 Rabbit and human eye responses after exposure to either 100 µL (Draize protocol) or 10 µL (LVET protocol) for the liquid household cleaner. All scoring done by the traditional Draize scoring scale. Compiled from Ghassemi *et al.* (1993) 50

Table 4-5 Average Time-to-Clear (days) for ocular effects following accidental exposure in humans and in rabbit eye irritation tests (LVET and Draize test) to household and cleaning products (Freeberg, Hooker et al. 1986)..... 51

Table 4-6 EPA Eye irritation toxicity categories (EPA 2003) 52

Table 4-7 Criteria for Classification of rabbits according to the GHS classification system..... 53

Table 4-8 Criteria for Classification of Substance According to the GHS Classification System (Modified from UN 2003) 54

Table 4-9 Composition of the light duty liquid detergent from the Beckley 1965 study (Beckley 1965) 56

Table 4-10 Mean Draize scores for individual ocular tissues of six rabbits, six dogs and four monkeys (unflushed) or three animals each (flushed) after instillation of 100 µL of a Light Duty Liquid Detergent (Beckley 1965) 56

Table 4-11 Composition of the test materials from the Beckley 1969 study (Beckley 1969) 57

Table 4-12 Liquid Household Cleaner composition used in the Ghassemi *et al.* (1997) study 58

Table 4-13 Composition of the test materials from the Roggeband, *et al.* (2000) study 59

Table 4-14 Ocular responses of humans and rabbits to identical volumes (3 µL) of Concentrated Laundry Liquid. Modified from Roggeband, *et al* (2000). 60

Table 4-15 Ocular responses of humans and rabbits to identical volumes (1 µL) of Concentrated Dishwshing Liquid. Modified from Roggeband, *et a* (2000) 60

Table 5-1 Description of number of unique materials tested in each assay system with corresponding *in vivo* data. 66

Table 5-2 Results of 105 unique materials tested in the Cytosensor assay and the rabbit LVET assay. Four of the materials were tested twice in the LVET assay and have toxicity categories from both tests listed. 71

Table 5-3 Distribution of product categories originally submitted with both animal eye irritation data and Cytosensor data 75

Table 5-4 Summary of Cytosensor data from the CTFA Phase III study using toxicity classifications determined by both the Draize Rabbit Test and the Low Volume Eye Test for surfactant-containing materials (Gettings, Lordo et al. 1996) 77

Table 5-5 Summary of Cytosensor and *in vivo* data from the COLIPA study which includes average values (see footnotes) from MA and CellTox AB laboratories (Brantom, Bruner et al. 1997) 79

Table 5-6 EpiOcular data paired with the Draize test 80

Table 5-7 Distribution of product categories for EpiOcular data paired with the Draize test 81

Table 5-8 EpiOcular data paired with LVET data 81

Table 5-9 Distribution of product categories for EpiOcular data paired with the LVET test 81

Table 5-11 Distribution of materials conducted in the BCOP assay. 84

Table 6-1 Distribution of product categories originally submitted with both animal eye irritation data and Cytosensor *in vitro* data. 85

Table 6-2 Contingency table depicting the accuracy and predictivity of the CM assay for EPA toxicity categories (determined by positive responses in the LVET) using cut-off values of MRD_{50} > 80 mg/mL = IV, 80 mg/mL >MRD_{50} > 2 mg/mL = III, and MRD_{50} < 2 mg/mL = I. The model does not propose to differentiate between EPA Category I and II materials. The total number of materials is listed as 108 since the three materials with differing repeat animal scores were each scored twice. 88

Table 6-3 Prediction results for the CM assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 89

Table 6-4 Contingency table depicting the accuracy and predictivity of the CM assay for GHS toxicity categories (determined by positive responses in the LVET) using cut-off values of MRD_{50} ≥ 10 mg/mL = NI, 10 mg/mL >MRD_{50} ≥ 2 mg/mL = 2B, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify GHS Category 2A materials. The total number of materials is listed as 108 since the three materials with differing repeat animal scores were each scored twice. 91

Table 6-5 Number of discordant results (and percentages) for the CM assay and GHS toxicity categories. 91

Table 6-6 Distribution of EPA categories for the 17 materials from the CM database classified as acid or alkaline. 92

Table 6-7 Contingency table depicting the accuracy and predictivity of the CM assay for EPA toxicity categories (determined by positive responses in the LVET) of non-acidic, non-alkaline materials using cut-off values of MRD_{50} ≥ 80 mg/mL = IV, 80 mg/mL >MRD_{50} ≥ 2 mg/mL = III, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify EPA Category II materials. 93

Table 6-8 Contingency table depicting the accuracy and predictivity of the CM assay for GHS toxicity cate8gories (determined by positive responses in the LVET)

using cut-off values of MRD_{50} ≥ 10 mg/mL = NI, 10 mg/mL >MRD_{50} ≥ 2 mg/mL = 2B, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify GHS Category 2A materials..... 95

Table 6-9 Distribution of product categories originally submitted with both animal eye irritation data and CTFA Phase III *in vitro* data..... 96

Table 6-10 Contingency table presenting the accuracy and predictivity of the CM for EPA toxicity categories (LVET-determined) for the 25 surfactant-based personal care products in the CTFA Phase III study (Gettings, Lordo et al. 1996)..... 98

Table 6-11 Discordant results for the CTFA CM study and EPA toxicity categories (LVET-determined)..... 98

Table 6-12 Contingency table presenting the accuracy and predictivity of the CM for EPA toxicity categories(Draize-determined) for the 25 surfactant-based personal care products in the CTFA Phase III study (Gettings, Lordo et al. 1996)..... 98

Table 6-13 Discordant results for the CTFA CM study and EPA toxicity categories (Draize-determined)..... 99

Table 6-14 Distribution of product categories originally submitted with both animal eye irritation data and COLIPA *in vitro* data..... 99

Table 6-15 COLIPA surfactant and surfactant containing materials. Contingency table depicting the concordance and predictivity of the CM assay for GHS toxicity classifications when the cut-off values shown in Figure 6-6 are applied..... 102

Table 6-16 Discordant results for the COLIPA CM study and GHS toxicity categories..... 102

Table 6-17 COLIPA surfactant and surfactant containing materials - Contingency table depicting the concordance and predictivity of the CM assay for EPA toxicity classifications when the cut-off values shown in Figure 6-7 are applied..... 102

Table 6-18 Discordant results for the COLIPA CM study and EPA toxicity categories..... 103

Table 6-19 Distribution of product categories originally submitted with both animal eye irritation data (LVET) and EpiOcular data..... 106

Table 6-20 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = IV, and ET_{50} < 4 min = I. ET_{50} values ≥4 min and <70 min are predicted to be EPA III. The model does not propose to identify EPA Category II materials..... 107

Table 6-21 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).. 108

Table 6-22 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = ET_{50} values ≥4 min and <70 min are predicted to be EPA III IV, and ET_{50} < 4 min = I. ET_{50} values ≥4 min and <70 min are predicted to be EPA III. The model does not propose to identify EPA Category II materials. 110

Table 6-23 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 110

Table 6-24 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials..... 112

Table 6-25 Prediction results for the EO assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 112

Table 6-26 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials..... 114

Table 6-27 Prediction results for the EO assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 114

Table 6-28 Distribution of product categories originally submitted with both animal eye irritation data (Draize) and EpiOcular data. 115

Table 6-29 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the Draize test) using cut-off values of ET_{50} ≥ 70 min = IV, and ET_{50} < 4 min = I. The model does not propose to identify EPA Category II materials..... 116

Table 6-30 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 116

Table 6-31 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the Draize test) using

cut-off values of ET_{50} ≥ 70 min = IV, and ET_{50} < 4 min = I. The model does not propose to identify EPA Category II materials. 118

Table 6-32 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 118

Table 6-33 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials. 119

Table 6-34 Discordant results for the EpiOcular assay and GHS toxicity categories. 120

Table 6-35 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} > 70 min = NI and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials. 121

Table 6-36 Discordant results for the EpiOcular assay and GHS toxicity categories. 121

Table 6-37 Distribution of product categories originally submitted with both animal eye irritation data and BCOP *in vitro* data. 123

Table 6-38 Contingency table (based on Figure 6-17) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 35 = II, and BCOP *in vitro* score < 35 = III. Although the model does propose to identify EPA Category II materials, there are no Category II's in the data set to test the hypothesis. The model does not propose to identify Category IV materials. 125

Table 6-39 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 125

Table 6-40 Contingency table (based on Figure 6-18) depicting the accuracy and predictivity of the BCOP assay for EPA classification (determined by the Draize test) using cut-off values of *in vitro* score > 75 = I, 75 > BCOP *in vitro* score > 25 = II, and BCOP *in vitro* score < 25 = III. The model does not propose to identify Category IV materials. 128

Table 6-41 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 128

Table 6-42 Contingency table (based on Figure 6-19) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories (determined by the

Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Category NL materials. 129

Table 6-43 Prediction results for the BCOP assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 130

Table 6-44 Contingency table (based on a combination of the results from Figure 6-20 & 6-21) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 25 = II, and a BCOP *in vitro* score < 25 = III. The model does not propose to identify Category IV materials. 133

Table 6-45 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 133

Table 6-46 Contingency table (based on Figure 6-22 & 6-23) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Category NL materials. 135

Table 6-47 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses). 135

Table 6-48 Scoring chart for histologically apparent damage and proposed EPA and GHS toxicity category. 137

Table 6-49 Integration of histopathology results with BCOP *in vitro* scores to give final EPA toxicity category classification (based on prediction model of Figure 6-18). Test material code letters appear in Figure 6-24 & 6-25. 137

Table 6-50 Contingency table (based on Figure 6-24 & 6-25) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 25 = II, and a BCOP *in vitro* score < 25 = III, plus histopathological evaluation. The model does not propose to identify Category IV materials. 141

Table 6-51 Discordant results for the BCOP assay and EPA toxicity categories. 141

Table 6-52 Contingency table (based on Figure 6-26 & 6-27) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories

(determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Nonirritant materials..... 144

Table 6-53 Discordant results for the BCOP assay and GHS toxicity categories..... 144

Table 7-1 Description of the results reported for each variability study. 148

Table 7-2 Within-laboratory reproducibility of CM from archived data that was originally obtained at Microbiological Associates, Inc. for the EC/HO study (Balls, Botham et al. 1995). The protocol utilized the CM using Transwells and an 810 second exposure time. At least triplicate runs were performed. 150

Table 7-3 Distribution of product categories for the within-laboratory reproducibility of the CM..... 151

Table 7-4 Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived Microbiological Associates, Inc. data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-nine total materials were tested. 152

Table 7-5 Non-Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived Microbiological Associates, Inc. data created for the COLIPA study for non-surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-nine materials were tested. 153

Table 7-6 Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived CellTox AB data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-six materials were tested... 153

Table 7-7 Non-Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived CellTox AB data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-six materials were tested... 154

Table 7-8 Distribution of product categories for the within-laboratory reproducibility of the COLIPA study..... 154

Table 7-9 Surfactant materials - Comparison of the MRD_{50} values for testing conducted approximately 21 months apart..... 155

Table 7-10 Non-surfactant materials - Comparison of the MRD_{50} values for testing conducted approximately 21 months apart..... 155

Table 7-11 Distribution of product categories for the intralaboratory reproducibility of the CM..... 155

Table 7-12 Positive Control Data of SLS completed at IIVS..... 156

Table 7-13 Surfactant Materials - Between-laboratories reproducibility of CM results from EC/HO study..... 158

Table 7-14 Non-surfactant materials - Between-laboratories reproducibility of CM results from EC/HO study..... 159

Table 7-15 Distribution of product categories for the interlaboratory reproducibility of the EC/HO study. 160

Table 7-16 Surfactant Materials - Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study..... 161

Table 7-17 Surfactant based formulations and mixtures - Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study..... 162

Table 7-18 Non-Surfactants, ingredients, and mixtures – Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study..... 163

Table 7-19 Distribution of product categories for the interlaboratory reproducibility of the COLIPA study..... 163

Table 7-20 EpiOcular intralaboratory repeatability both within run and between experiments..... 164

Table 7-21 Distribution of product categories for the intralaboratory repeatability of the EpiOcular assay. 166

Table 7-22 Intralaboratory reproducibility of EpiOcular tissue over a nine year period from 1997 through 2005 for two different laboratories..... 166

Table 7-23 Standard deviation range for 0.3% Triton X-100 for EpiOcular tissue over a nine year period..... 166

Table 7-24 Interlaboratory reproducibility of four laboratories in the Colgate-Palmolive Phase II validation study. 168

Table 7-25 Interlaboratory reproducibility of two laboratories in the Colgate-Palmolive Phase III validation study..... 168

Table 7-26 Distribution of product categories for the interlaboratory reproducibility of the EpiOcular assay. 170

Table 7-27 BCOP within run reproducibility..... 172

Table 7-28 Distribution of product categories for the within-run reproducibility of the BCOP assay. Some products have repeat tests. 189

Table 7-29 Intralaboratory reproducibility for 5 antimicrobial cleaning products. See Table 7-27 for individual cornea scores..... 191

Table 7-30 Distribution of product categories for the intralaboratory reproducibility for antimicrobial cleaning products..... 192

Table 7-31 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Gautheron *et al.* (1994)[1]..... 193

Table 7-32 Distribution (estimated) of product categories for the interlaboratory reproducibility for the Gautheron study..... 194

Table 7-33 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Balls *et al.* (1995) 195

Table 7-34 Distribution of product categories (estimated) for the interlaboratory reproducibility for the Balls study..... 196

Table 7-35 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Southee (1998) 197

Table 7-36 Distribution of product categories (estimated) for the interlaboratory reproducibility for the Southee study..... 197

List of Figures

Figure I The proposed testing strategy for evaluating the EPA toxicity category for anti-microbial cleaning products. xxiv

Figure 1-1 The proposed testing strategy for determining the EPA toxicity category for anti-microbial cleaning products. 6

Figure 1-2 Factors that impact exposure to the eye 12

Figure 1-3 Summary of the Depth of Injury Model 13

Figure 1-4 Example of the metabolic rate data as a function of surfactant type and concentration 14

Figure 1-5 Photomicrographs of a) the EpiOcular model showing the stratification and lack of surface keratinization (photo from MatTek Corporation, Ashland, MA), b) the cornea of a rabbit eye (photo courtesy of MatTek Corporation, Ashland, MA), and c) a human cornea. 17

Figure 1-6 A cross-section of a typical bovine cornea as used in the BCOP assay. (H&E stain) 20

Figure 2-1 Diagram of the operating components of the silicon microphysiometer (Bruner, Miller et al. 1991) 29

Figure 2-2 The original silicon microphysiometer sensor chamber with the coverslip in place (Bruner, Miller et al. 1991) 29

Figure 2-3 Diagram of the operating components of the Cytosensor (Cytosensor Manual) 29

Figure 2-4 The Cytosensor chamber with the Transwell in place (Cytosensor Manual) 30

Figure 2-5 A comparison of data obtained from 11 surfactant-containing products with SM and CM. 33

Figure 2-6 Diagrammatic representation of EpiOcular tissue growing in a milliicell chamber placed within a well of a 24-well plate. A photomicrograph of a cross section through the tissue and underlying membrane is included. 34

Figure 2-7 Diagrammatic representation of the testing procedure using EpiOcular tissue. Incubation is carried out at 37°C, and test material is thoroughly removed before the addition of MTT. 36

Figure 2-8 Photographs of various aspects of the EpiOcular assay. 37

Figure 2-9 Photographs of various procedures occurring in the BCOP protocol. Upper left – Placing an excised cornea on the corneal holder. Upper right – Using the opacitometer to measure the opacity of a bovine cornea contained in a corneal holder. Bottom left – Visual comparison of the transparency of an untreated cornea on the left and a cornea treated with an irritating material on the right. Lower right – removing fluorescein solution from the posterior chamber prior to measuring its optical density in a spectrophotometer. 38

Figure 2-10 Histological evaluation of corneas. 41

Figure 4-1. Performance of the Silicon Microphysiometer in predicting the Draize MAS score for test materials from the CTFA Phase III study of surfactant-based formulations (Gettings, Lordo et al. 1996). The variability associated with both the animal test and the *in vitro* test is shown on the graph. 61

Figure 6-1 Cytosensor MRD_{50} values plotted against EPA toxicity categories determined by the LVET. Suggested cut-off values with their predicted EPA categories are included. There are 105 unique materials; however, 3 materials are graphed with 2 different EPA categories since they were tested twice in the animal trials with different results each time. 87

Figure 6-2 Cytosensor MRD_{50} values plotted against GHS toxicity categories determined by the LVET. All materials except oxidizing formulations are graphed. Suggested cut-off values with their predicted GHS categories are included. There are 105 unique materials; however, 3 materials have 2 GHS categories each since they were tested twice in the animal trials. 90

Figure 6-3 Cytosensor MRD_{50} values plotted against EPA toxicity categories determined by the LVET. Only non-acidic, non-alkaline materials are graphed. Suggested cut-off values with their predicted EPA categories are included. There are 100 unique materials; however, 3 materials have 2 values since they were tested twice in the animal trials. 92

Figure 6-4 Cytosensor MRD_{50} values plotted against GHS toxicity categories determined by the LVET. Only non-acidic, non-alkaline materials are graphed. Suggested cut-off values with their predicted GHS categories are included. There are 100 unique materials; however, 3 materials have 2 values since they were tested twice in the animal trials. 94

Figure 6-5 Plot of CM data versus both LVET- and Draize-defined EPA Categories for the 25 surfactant-based personal care products tested in the CTFA Phase III (Gettings, Lordo et al. 1996) evaluation using cut-off values of $MRD_{50} \geq 80$ mg/mL = IV, 80 mg/mL $> MRD_{50} \geq 2$ mg/mL = III, and $MRD_{50} < 2$ mg/mL = I. The model does not propose to identify EPA Category II materials. 97

Figure 6-6 Surfactant and surfactant-containing formulation results of the COLIPA study related to GHS classification. Data points indicate the mean MRD_{50} for both laboratories (with the exception of two data points where only one laboratory made the determination). In some cases data points have been slightly offset along the X-axis in order to clearly separate them from data of similar magnitude. 100

Figure 6-7 Surfactant and surfactant-containing formulation results of the COLIPA study related to EPA classification. Data points indicate the mean MRD_{50} for both laboratories with the exception of 24 and 52 which were done in one laboratory only. In some cases data points have been slightly offset along the X-axis in order to clearly separate them from data of similar magnitude. The individual materials can be identified by comparing the numbers adjacent to the symbols with the numbering code given in Table 5.3.1.3. 101

Figure 6-8 Distribution of CM scores for the products without animal data using cut-offs of $MRD_{50} \geq 80$ mg/mL = IV, 80 mg/mL $> MRD_{50} \geq 2$ mg/mL = III, and $MRD_{50} < 2$ mg/mL = I. 103

Figure 6-9 EpiOcular ET_{50} values plotted against EPA categories determined by the LVET. Suggested cut-off values with their predicted EPA categories are included..... 107

Figure 6-10 EpiOcular ET_{50} values plotted against EPA categories determined by the LVET. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted EPA categories are included..... 109

Figure 6-11 EpiOcular ET_{50} values plotted against GHS categories determined by the LVET. Suggested cut-off values with their predicted GHS categories are included..... 111

Figure 6-12 EpiOcular ET_{50} values plotted against GHS categories determined by the LVET. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted GHS categories are included..... 113

Figure 6-13 EpiOcular ET_{50} values plotted against EPA categories determined by the Draize test. Suggested cut-off values with their predicted EPA categories are included..... 115

Figure 6-14 EpiOcular ET_{50} values plotted against EPA categories determined by the Draize test. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted EPA categories are included..... 117

Figure 6-15 EpiOcular ET_{50} values plotted against GHS categories determined by the Draize. Suggested cut-off values with their predicted GHS categories are included. 119

Figure 6-16 EpiOcular ET_{50} values plotted against GHS categories determined by the Draize. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted GHS categories are included. 120

Figure 6-17 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. 124

Figure 6-18 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA toxicity categories for test materials BR and BS were determined by using the results of an LVET assay. The discussion of the materials labeled as “High solvent” occurs later in this chapter. 127

Figure 6-19 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories for test materials BR and BS were determined by using the results of an LVET assay. The discussion of the materials labeled as “High solvent” occurs later in this chapter. 129

Figure 6-20 BCOP *in vitro* scores (3 minute exposure) for High Solvents are plotted against EPA categories determined by the Draize test. Five High Solvent materials had 10 minute data only and therefore are not included in this graph. Proposed cut-off values with their predicted EPA categories are included. 131

Figure 6-21 BCOP *in vitro* scores for non-High Solvent materials plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA categories for test materials BR and BS were determined by using the results of an LVET assay. 132

Figure 6-22 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Five materials had only 10 minute data and therefore are not included on this graph. Proposed cut-off values with their predicted GHS categories are included. Test material BB is not included due to the study criteria not being met for the GHS category. 134

Figure 6-23 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay. 134

Figure 6-24 BCOP *in vitro* scores (3 minute exposure) for High Solvent formulations plotted against EPA categories determined by the Draize test. Five High Solvent materials had 10 minute data only and therefore are not included in this graph. Materials with histology-determined EPA categories are circled with the final category indicated. 139

Figure 6-25 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay. 140

Figure 6-26 BCOP *in vitro* scores (3 minute exposure) for High Solvent formulations plotted against GHS categories determined by the Draize test. Five High Solvent materials had only 10 minute data and therefore are not included on this graph. Proposed cut-off values with their predicted GHS categories are included. Materials with histology-determined EPA categories are circled with the final category indicated. Test material BB is not included due to the study criteria not being met for the GHS category. 142

Figure 6-27 BCOP *in vitro* scores for non-High solvent materials plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay. Materials with histology-determined EPA categories are circled with the final category indicated. 143

Figure 7-1 Graph of 10% SLS (positive control) MRD_{50} values obtained at IIVS over a 28-month period. 157

Figure 11-1 Process of safety evaluations 202

Annexes

ANNEX A (Protocols) A1
A1 IIVS Cytosensor Protocol A2
A2 COLIPA Cytosensor Protocol A12
A3 IIVS EpiOcular Protocol A23
A4 IIVS BCOP Protocol A33

ANNEX B (Formulations and Physicochemical Properties of Formulations) B1
B1 CHEMICALS B2
B2 FORMULATIONS B12
B3 Formulations for Cytosensor Data Paired with LVET Data B31
B4 CTFA Phase III Formulations B48
B5 COLIPA Formulations B54

ANNEX C (Animal Data) C1
C1 EpiOcular and BCOP Animal Data C2
C2 Cytosensor Animal Data C67
C3 CTFA Animal Data C68
C4 CTFA LVET Animal Data C94
C5 CTFA Draize Animal Data C120
C6 CTFA LVET Animal Data– 3 Day Averages C146
C7 COLIPA Animal Data C151
C8 COLIPA Draize Animal Data C173
C9 Company Paired LVET Animal Data C179
C10 Company Paired LVET Animal Raw Data C260

ANNEX D (BCOP In Vitro Data) D1

ANNEX E (Cytosensor *In Vitro* Data) E1
E1 IIVS Positive Control Data E2
E2 COLIPA Raw Data from CellTox AB E10
E3 COLIPA Raw Data from Microbiological Associates, Inc. E16
E4 COLIPA Decoded Samples for Microbiological Associates, Inc. E20

ANNEX F (BCOP BRD) F1

ANNEX G (DRAFT BCOP HISTOPATHOLOGY GUIDANCE DOCUMENT) G1

List of Abbreviations

AC	Acidic (used to designate a formulation “bucket”)
AISE	European Soap and Detergent Industry Association
AL	Alkaline (used to designate a formulation “bucket”)
BCOP	Bovine Corneal Opacity and Permeability Assay
BRD	Background Review Document
CM	Cytosensor Microphysiometer
COLIPA	European Cosmetic, Toiletry, and Perfumery Association
CPSC	Consumer Products Safety Commission
CTFA	U.S. Cosmetics, Toiletries, and Fragrance Association
CV	Coefficient of Variance
DPIC	Drug & Poisons Information Centre
DMEM	Dulbecco’s Modified Eagle’s Medium
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
EC/HO	European Commission/British Home Office
EO	EpiOcular™
EPA	Environmental Protection Agency
EU	European Union
FHSA	Federal Hazardous Substances Act
FIFRA	Federal Insecticide, Fungicide, and Rodenticide Act
GHS	United Nations Globally Harmonized Systems
ICCVAM	Interagency Coordinating Committee on the Validation of Alternative Methods
IIVS	Institute for In Vitro Sciences, Inc.
IRAG	Interagency Regulatory Alternatives Group
LVET	Low Volume Eye Test
MA	Microbiological Associates, Inc.
MMAS	Modified Maximum Average Score
MRD_{50}	Metabolic rate decrement of 50%
MTT	3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
NICETAM	National Toxicology Program (NTP) Interagency Center for the Evaluation of Alternative Toxicological Methods
OPP	Office of Pesticide Programs
P&G	The Procter & Gamble Company
PBS	Phosphate Buffered Saline
pH	An acidity/alkalinity index; the logarithm of reciprocal of the hydrogen ion concentration
RC	Reactive chemistry, more generally referred to in this BRD as oxidizer (used to designate a formulation “bucket”)
SD	Standard Deviation
SEM	Standard Error of the Mean
SLS	Sodium Lauryl Sulfate
SM	Silicon Microphysiometer
SO	Solvent (used to designate a formulation “bucket”)
SU	Surfactant (used to designate a formulation “bucket”)
TSCA	Toxic Substances Control Act

Acknowledgements

The generous contributions from the following companies and individuals who provided *in vitro* and/or *in vivo* data considered in this Background Review Document are gratefully acknowledged:

Clorox , Pleasanton, CA USA
Colgate-Palmolive Company, Piscataway, NJ USA
The Dial Corporation, Scottsdale, AZ USA
EcoLabs, St. Paul, MN USA
JohnsonDiversey, Inc., Sturtevant, WI USA
S.C. Johnson & Son, Inc., Racine, WI USA
The Procter and Gamble Company, Cincinnati, OH USA

Institute for In Vitro Sciences, Inc., Gaithersburg, MD USA

Greg Mun and Hans Raabe (Institute for In Vitro Sciences, Inc.)

Preface

On June 4, 2004, Mr. James Jones, Director, Office of Pesticide Programs, EPA informed Dr. William Stokes, Director, ICCVAM that the EPA was developing, via a subgroup of the Pesticide Program Dialogue Committee, a non-animal assessment approach for evaluating the eye irritation potential of antimicrobial cleaning products for the purpose of determining appropriate product cautionary labeling. Mr. Jones requested that ICCVAM conduct a technical review of this approach when finalized.

This approach has been finalized and is presented in Figure I as a flowchart which outlines how the EpiOcular (EO) Assay, Cytosensor Microphysiometer (CM) Assay and Bovine Corneal Opacity and Permeability (BCOP) Assay are to be used to determine the EPA toxicity Category (I – IV) with regards to ocular cautionary labeling for anti-microbial cleaning products.

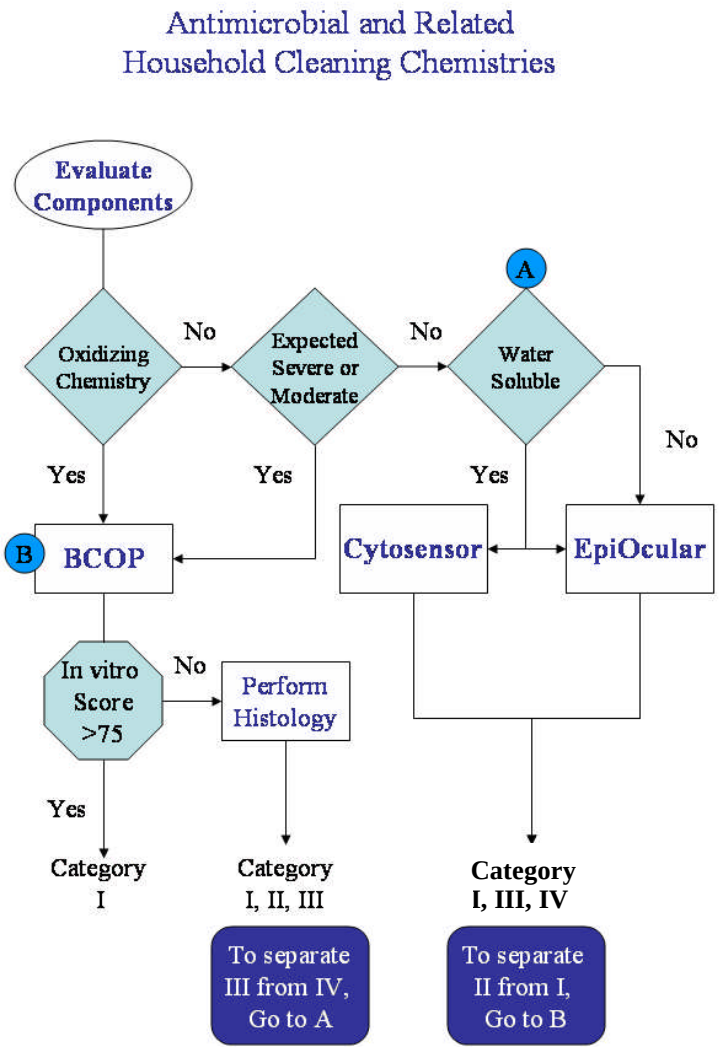


Figure I The proposed testing strategy for evaluating the EPA toxicity category for anti-microbial cleaning products.

Based on the request of Mr. Jones, we now ask ICCVAM to conduct a technical review of the attached approach and supporting materials and develop an opinion on whether use of this approach will assure the EPA that, with a reasonable level of certainty, no antimicrobial product will be underlabeled.

Executive Summary

This Background Review Document (BRD) presents a description of an *in vitro* testing strategy for determining the appropriate product cautionary labeling for anti-microbial cleaning products. The strategy is flexible in that several different assays can be used either alone or combined with a second assay to obtain an EPA or GHS toxicity category. The three assays proposed are the Cytosensor Microphysiometer (CM) assay, the EpiOcular™ (EO) assay (MatTek Corporation, Ashland MA), and the Bovine Corneal Opacity and Permeability (BCOP) assay. A complete description of these assays and data supporting their predictive capacity and reproducibility are contained in the BRD.

This BRD is a joint project of seven companies – The Clorox Company, Colgate-Palmolive, The Dial Corporation, EcoLabs, JohnsonDiversey, Inc., S.C. Johnson & Son, Inc. and The Procter & Gamble Company - who manufacture anti-microbial cleaning products. Normally cleaning products are regulated by the US Consumer Product Safety Commission (CPSC), but when the product is labeled as “anti-microbial” – it is then classified as a pesticide and falls under the jurisdiction of the EPA. Registration of such products requires animal testing for several endpoints, including eye irritation, to determine the appropriate product cautionary labeling. Since many products of this type have been safely marketed (minus the anti-microbial claim) without animal testing, the companies wished to provide data supporting the position that *in vitro* test methods for eye irritation could provide adequate cautionary labeling.

The companies therefore provided the animal eye irritation data (using both the standard Draize test and the Low Volume Eye Test [LVET]) that were available in their files for a large set of cleaning products. At the same time data from one or more of the *in vitro* tests listed above was provided for each material, or was newly generated. These paired data sets were used to determine the predictive ability of the three *in vitro* methods. In addition, the within laboratory and between laboratory reproducibility of the *in vitro* methods was assessed.

As a guideline against which to asses the performance of the *in vitro* methods, an analysis of the reproducibility of the rabbit eye test was presented which shows that this *in vivo* method does not always give the same EPA toxicity category when multiple tests are run. Thus the *in vitro* methods should not be expected to provide a 100% duplication of the animal results.

In addition to data provided by the participating companies for the anti-microbial cleaning products, other historical studies which were conducted with similar ingredients (e.g. surfactants) or mixtures are also presented and analyzed.

It was found that each of the three *in vitro* tests had different areas of strength. The CM and EO assays were more sensitive and thus are useful to separate EPA category III materials from EPA category IV materials. These materials are in the milder

side of the toxicity range. In contrast, the BCOP assay uses a more robust tissue and therefore is able to differentiate between EPA category I materials and EPA category II materials. These materials are in the higher side of the toxicity range. A diagram of this strategy is presented in the Preface and in Section 1. Introduction and Rationale for the Proposed Test Method.

The proposed *in vitro* strategy is very conservative and results in over labeling of some products, especially many EPA category IV materials which are overpredicted to be EPA category III. The participating companies are aware of these overpredictions and have accepted it as a small consequence of adopting non-animal testing strategy.

Test Method Predictive Capacity

Prediction models for the three *in vitro* assays (CM, EO and BCOP) were constructed using the same approach (a graphical one). For each model all the paired *in vitro* and *in vivo* data provided were used, and the *in vitro* data were plotted against the *in vivo*-defined toxicity category (both EPA and GHS). In some cases only data from an LVET assay were available, and in other cases only data from a Draize test were available. Generally each type of data was analyzed separately, although it was concluded that the prediction models were the same regardless of the *in vivo* assay used.

Once the data were graphed, cut-off lines were fitted by eye to provide the “best” predictions. A description of these cut-offs then became the prediction model. The strategy in setting the cut-offs was to minimize under predictions of toxicity at the expense of over predictions. Of course, over and under predictions are somewhat arbitrary terms since we have shown earlier in this BRD (Section 4.8.1) that repeated three-rabbit eye irritation tests do not necessarily provide identical toxicity classifications. In other words, a second rabbit test may over or under predict the first test.

Although data from the testing of anti-microbial cleaning products (and related cleaning products) were primarily used to set the cut-offs, additional data from chemically related formulations and some pure substances (e.g. surfactants) were used to provide supporting information for our decisions.

Summary contingency tables showing concordance, under prediction and over prediction are presented below for each of the methods.

a) Cytosensor

The following table shows the performance of the Cytosensor in predicting the EPA toxicity category (defined by the LVET test) of 108 cleaning products. There were no underpredictions of EPA toxicity categories, but 89% of the Category IV materials were overpredicted as Category III or higher. However

the CM was able to clearly identify some Category IV materials. Results for the prediction of GHS categories were similar.

The CM should be useful in clearly identifying materials as EPA Category III or Category IV, but cannot separate EPA toxicity category I from category II. Oxidizing materials, or materials not completely aqueous soluble at the highest dilution, should not be tested in the CM.

LVET- Determined EPA Category	CM Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	9	0	0	9	100%	NA	0%
II	11	0	0	11	0%	100%	0%
III	40	20	0	60	33%	67%	0%
IV	4	21	3	28	11%	89%	NA
Total	64	41	3	108	30%		
Predictivity	14%	49%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	86%	51%	NA				

b) EpiOcular™

Animal eye irritation data from both the Draize test and the LVET were supplied paired with EO data. The following two tables show the performance of the EpiOcular™ assay in predicting the EPA toxicity categories defined by the by each of the *in vivo* tests. There was only one underprediction for the 41 total materials. The EO method was able to clearly separate a few EPA category IV materials, although most Category IV materials will be overpredicted as Category III. Results for the prediction of GHS categories were similar.

The EO assay should be useful in clearly identifying materials as EPA Category III or Category IV, but cannot separate EPA toxicity category I from category II. Oxidizing materials should not be tested in the CM, but both water soluble and water insoluble materials can be tested.

LVET- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	1	0	0	1	100%	NA	0%
II	0	0	0	0	0%	0%	0%
III	2	7	0	9	78%	22%	0%
IV	2	4	0	6	0%	100%	NA
Total	5	11	0	16	50%		
Predictivity	20%	64%	0%				
Category under predicted	NA	0%	0%				
Category over predicted	80%	36%	NA				

Draize- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	12	0	0	12	100%	NA	0%
II	0	1	0	1	0%	0%	100%
III	1	3	0	4	75%	25%	0%
IV	1	4	3	8	38%	63%	NA
Total	14	8	3	25	72%		
Predictivity	86%	38%	100%				
Category under predicted	NA	12%	0%				
Category over predicted	14%	50%	NA				

c) BCOP

The vast majority of animal data used in the analysis of the BCOP assay were from the Draize test; only two tests were conducted using the LVET. Histopathological examination of the treated bovine corneas was included in the analysis in addition to the traditional *in Vitro* Score which measures the opacity and permeability of the cornea.

The following table shows the performance of the BCOP assay (including histopathology) in predicting EPA toxicity categories. Only 2 of 61 materials (8%) were underpredicted. All of the EPA toxicity category IV materials are overpredicted as Category III since the BCOP does not seem to be able to differentiate between materials at this lower end of the toxicity scale. The BCOP assay does differentiate between EPA Category I and II materials, so it is most useful in this higher range.

If the anti-microbial cleaning product is a High Solvent (>5 solvent) formulation, it should be tested in the BCOP assay using a 3 minute exposure instead of the normal 10 minute exposure.

Draize- Determined EPA Category	BCOP Predicted (with histology) EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	II	III	Total			
I	23	2	0	25	92%	NA	8%
II	4	1	0	5	20%	80%	0%
III	3	2	7	12	58%	42%	0%
IV	0	1	18	19	0%	100%	NA
Total	30	6	25	61	51%		
Predictivity	77%	17%	28%				
Category under predicted	NA	33%	0%				
Category over predicted	23%	50%	72%				

<u>Test Method Reliability</u>

This parameter was assessed by measuring the within and between laboratory reproducibility for each of the *in vitro* methods. Within assay repeatability was also assessed when the values were available. The coefficient of variation (CV) between repeat values was used as a measure of reliability.

Although the primary data used to calculate the CV's was from studies with anti-microbial (or similar cleaning products), the BRD also contains supporting data from other studies which used individual ingredients or mixtures (e.g. of surfactants).

a) Cytosensor

Within laboratory reproducibility was assessed from the results of two international validation studies. In the first study (EC/HO study), the mean CV for 31 chemicals (three CM runs each) was 23.9%. For the second study (Colipa eye irritation validation), one laboratory had a mean CV of 19.7% for surfactant materials and 15.4% for non-surfactant materials. A second laboratory had a mean CV of 14.3% for the surfactant materials and 10.4% for the non-surfactant materials.

Interlaboratory reproducibility was also assessed from data generated in the above validation studies. In the HO/EC study, four laboratories had a mean between laboratory CV of 37% for surfactant materials and 50.6% for non-surfactant materials. For the Colipa study two laboratories had a mean between laboratory CV of 23.3% for surfactant materials, 16.5% for surfactant-based formulations and mixtures, and 32.5% for non-surfactant ingredients and mixtures.

b) EpiOcular™

Within laboratory reproducibility was estimated from the repeated testing of a single material (0.3% Triton X-100) over a nine year period in two laboratories. The CV for these repeats was 20.7%.

Interlaboratory reproducibility was assessed from two phases of a validation study conducted by Colgate-Palmolive. Nineteen pure surfactants and mixtures were tested by four laboratories in Phase I with a mean between laboratories CV of 18.1%. Fifty-four pure surfactants and mixtures were tested by two laboratories in Phase II with a mean between laboratories CV of 11.8%.

c). BCOP

Within run reproducibility was estimated for the BCOP assay from anti-microbial cleaning products tested for this BRD. When the overall *In Vitro Score* was low (≤10), the within run CV could be quite high (mean CV = 266% for opacity and 167% for permeability) because small changes in low numbers result in high CV's. However, such small differences in magnitude in opacity or permeability scores are relatively meaningless with respect to the overall

range of scores that is possible. However for materials where the mean *In Vitro Score* was >10, the mean CV for opacity was 27.9% and for permeability was 24.1%.

c) BCOP

BCOP Intralaboratory reproducibility for the anti-microbial cleaning products was 20.3% for five materials (2 – 6 values per material). Intralaboratory CV's found by NICEATM in their BCOP Test Method Review Document ranged from 12.6% to 14.8%.

Interlaboratory reproducibility for the BCOP assay was assessed from three studies where the median CV's were: Study 1 (11-12 laboratories) 46.9%, Study 2 (5 laboratories) 30.6% and Study 3 (3 laboratories) 22.8%. The median CV is presented for these studies since the mean CV was strongly affected by large CV's for materials where the overall *In Vitro Score* was ≤10.

Overall Testing Strategy

A strategy is presented in this BRD where materials can be tested in one or more *in vitro* assays to reach a final EPA or GHS toxicity category. Oxidizing formulations are always tested in the BCOP assay, but other formulation types could be tested in any of the three assays, as long as their physical characteristics are compatible with that system. However a second assay may be needed since the BCOP can not identify an EPA category IV material, while the CM and EO are able to. Conversely the BCOP assay may be used differentiate between an EPA toxicity category I and II, but the CM and EO are not able to do that.

1 Introduction and Rationale for the Proposed Test Method

1.1 Introduction

1.1.1 Description of framework for development of program

For the past twenty years, extensive research has been conducted to develop non-animal approaches for evaluating the eye irritation potential of household and commercial cleaning products. This research involved developing a detailed understanding of the mechanism by which these products induced eye injury and then developing *in vitro* and *ex vivo* assays that modeled that mechanism.

In the mid to late 1990's, manufacturers of household and commercial cleaning products started to conduct internal evaluations of these assays to evaluate whether they could be used to determine the appropriate ocular precautionary labeling for their specific products.

These internal studies were successful and for nearly a decade these non-animal methods together with a weight-of-evidence approach have been used in lieu of traditional rabbit models for the determination of ocular precautionary labeling of products.

For the vast majority of household and commercial cleaning products, the Consumer Products Safety Commission (CPSC) has regulatory authority for ocular cautionary labeling. The CPSC itself actively encouraged companies to use non-animal tests. Its publication "Requirements under the Federal Toxic Substances Act: Labeling and Banning Requirements for Chemicals and Other Toxic Substances" states:

> "The FHSA only requires that a product be labeled to reflect the toxicities it presents. It does not require anyone to perform animal tests. The Commission policy is, whenever possible, to evaluate product toxicities by using alternatives to animal testing. We encourage anyone evaluating products to determine whether they present toxicities listed in the FHSA to follow a similar policy."

A small percentage of household and commercial cleaning products carry the claim, "anti-microbial". These are considered pesticidal products and regulatory authority for ocular precautionary labeling for these products rests with EPA's Office of Pesticide Programs (OPP). <u>In contrast to regulations for non-pesticidal cleaning products, EPA regulations for pesticide registration require that animal tests be performed to determine ocular precautionary labeling.</u>

Since non-animal methods are predominately used today to determine the ocular precautionary labeling for the vast majority of household and commercial cleaning products, a project (which has resulted in this Background Review Document) was

initiated with the goal of gaining adoption of these methods for ocular precautionary labeling decisions for a subset of specific products regulated by OPP – *i.e.,* anti-microbial cleaning products.

Within this document is a proposed approach and supporting materials which outline how these non-animal methods can be used to determine the EPA toxicity Category (I – IV) for ocular cautionary labeling of anti-microbial cleaning products.

It is now requested that ICCVAM conduct a technical review of this approach and supporting materials and develop an opinion on whether the use of this approach will assure the EPA that, with a reasonable level of certainty, antimicrobial cleaning products will not be under labeled.

1.1.2 Summary of Project History

The genesis of the herein described non-animal testing approach occurred within the Pesticide Program Dialog Committee, a Federal Advisory Committee established to advise EPA on the concerns of its many and diverse stakeholders. The concern was broached in this committee that since cleaning products had apparently been safely marketed for many years without the use of new animal tests, it seemed unreasonable to force them to be tested in animals just because of a different claim. Their thought was that as long as the non-animal methods would allow products to be adequately labeled, then those options should be available and acceptable.

EPA/OPP Director Jim Jones agreed with the advice of the committee to investigate the feasibility of accepting non-animal methods for the labeling of cleaning products, and began supporting efforts to develop a non-animal testing approach. The effort was taken up by two major manufacturers of anti-microbial cleaning products, the Procter & Gamble Company and S.C. Johnson & Son, Inc. A specialized *in vitro* laboratory – The Institute for *In Vitro* Sciences, Inc. (IIVS) was asked to help coordinate the program, perform any needed testing, and prepare the eventual submission.

Although the project was originally scheduled to be presented directly to the EPA's science advisory panel, it was later determined that the Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) would oversee the technical review and then present their findings and recommendations to the EPA. Therefore, this submission is being prepared according to the formatting suggested by ICCVAM.

To initiate the project, companies that manufacture anti-microbial cleaning products or materials with similar formulations were invited to participate and to share their animal data, *in vitro* data, and toxicological expertise. If this program is successful, there will be several advantages for a manufacturer, for example, the ability to:

- normalize standard practices for non-regulated product development with regulated product requirements, and

- use formulation development data obtained *in vitro* to support registration and labeling

The following seven companies agreed to assist the project by supplying animal and/or *in vitro* data:

- Clorox
- Colgate -Palmolive Company
- The Dial Corporation
- EcoLabs
- JohnsonDiversey, Inc.
- S.C. Johnson & Son, Inc.
- The Procter & Gamble Company

Each company was informed that the specific data that they contributed would be coded so that it could not be linked directly to them. They were asked to supply the following type of information for each cleaning formulation that would be used in the program:

1) Complete data (carried out to 21 days) from individual animals used to test a substance
2) Detailed description of the animal test protocol, if possible
3) Characterization of the suspected chemical activity category of the formulation (see below)
4) Description of the ingredients contained in the test formulation at the level of detail that would be supplied to a poison control center
5) Description of the *in vitro* test used with the test substance
6) Raw data from the *in vitro* test, if possible

A sample Excel® spread sheet was provided to each potential participant which included the input form that each submitter was asked to fill out for each animal tested with each formulation.

The following chemical descriptors were suggested to characterize the different types of chemically-induced mechanisms associated with ocular irritation. These were chosen based on existing information about the mechanisms of ocular irritation and the common types of formulation chemistries used in commercial and household cleaning products.

- Surfactants (SU) (e*.g.,* cationic, anionic, and nonionic with limited acid or alkaline activity)
- Acids (AC) (*e.g.,* with pH <4, especially where reserve acidity would contribute to the irritation potential)
- Alkaline (AL) products (bases) (*e.g.,* with pH >9, especially where reserve alkalinity would contribute to the irritation potential)

- Solvents (SO) (where organic solvents are expected to contribute to the irritancy potential (*e.g.,* alcohols, glycol ethers, etc.))
- Oxidizers (RC; reactive chemistry) (formulations containing specific reactive chemicals, *e.g.,* hypochlorite, peroxide, percarbonate, oxygen bleaches, etc.)

The process began by collecting data (both animal and non-animal data) from the historic records of the participating companies and combining it in a database (at IIVS) to determine the effectiveness of the methods to predict the EPA toxicity labeling categories of anti-microbial products. We compared the specific EPA categories with the *in vitro* scores to determine prediction models for each *in vitro* test which could be used to set cut offs for the various categories. Since knowing the correct EPA toxicity category for the substances was imperative, raw data for the individual test animals were absolutely required.

In vivo methods: Data from two types of rabbit tests were submitted during this project. One set was from the traditional Draize rabbit eye test, and the second was from a similar test – the Low Volume Eye Test (LVET). The LVET is also a rabbit eye test, but it differs from the traditional Draize assay in the volume tested and the location on which the material is placed on the eye. The LVET uses one-tenth the volume of the Draize test (10 µL vs. 100 µL) and places the material directly on the central surface of the cornea as opposed to instilling the material in the conjunctival sac. This volume and placement is thought to more closely mimic a typical human accidental exposure. Excel spreadsheets were created to convert raw animal data into the appropriate EPA or GHS scoring scale.

In vitro methods: Three different *in vitro* assays for eye irritation were in common use by the participating manufacturers. These were the EpiOcular (EO) assay, the Cytosensor Microphysiometer (CM) assay, and the Bovine Cornea Opacity and Permeability (BCOP) assay. The EO assay is a three-dimensional, non-keratinized, tissue constructed from human epithelial cells. It is designed to have a similar construction and histological appearance to the epithelial cell layers covering the cornea. The CM is an instrument which measures changes in the metabolism of cells. Increasing amounts of test article are exposed to the cells until the metabolic rate falls by 50% (MRD_{50}). The lower the MRD_{50} value, the higher is the potential for eye irritation. The BCOP assay uses isolated bovine corneas dissected from whole globe eyes obtained from slaughterhouses. Test substances can be placed directly on the surface of these corneas and subsequent changes in both the opacity and the barrier function of the epithelial cell layer can be measured. Additionally, histopathology can be performed on the corneas so that the induced damage can be visualized.

Anti-microbial cleaning products can be formulated in different ways. To prepare for the possibility that each different type of formulation might have a slightly different pattern of toxicity when used in the different *in vitro* tests, we described each product according to what was thought would be the major driver of eye irritation for that product. The descriptors chosen were acid, alkaline, oxidizer, surfactant, and solvent.

The first part of our study was purely retrospective. Graphical comparisons between the toxicity categories determined by the *in vivo* and the *in vitro* scores for the same test materials were made. This helped to decide if sufficient materials were available in each toxicity category to allow the determination of potential cut-off values that would ultimately define EPA classifications. Although this determination was possible in some cases, in others we found that the data were lacking to clearly indicate where the cut-off values should fall. However, probable cut-off ranges were still hypothesized based on the distribution of the data and known irritation profiles determined based on *in vivo* animal data.

Materials were then sought with which to generate additional *in vitro* data from the database of animal studies without paired *in vitro* data. It was hoped that these additional studies would clarify where the cut-off values should lie. Attempts were made to find materials from the toxicity categories that had low representation (for example, EPA Category II materials were significantly underrepresented), or where the cut-off values were difficult to determine. These materials were requested from the appropriate manufacturers, and if the manufacturers chose to have them tested in the *in vitro* assay that was suggested, they were instructed to code the materials before submitting them for testing. The materials were then tested under code at IIVS. If the *in vitro* test selected was the BCOP assay, the corneas were also submitted for histopathology which was conducted either by IIVS staff or by an IIVS contractor skilled in ocular histopathology. The histopathology results were then compared to the BCOP *in vitro* scores to determine if they were reflective of the *in vitro* scores, or if the toxicity category of the material should be increased. Materials were not decoded until after the final decision as to the ocular irritation potential of the substance was made.

These new data were then combined with the previous data to determine if they supported the initial determination of cut-offs or if they provided more information which allowed a better estimation of the cut-off.

After the predictive capacity of each *in vitro* test was examined, we investigated whether any of the tests could be stand-alone predictors of all of the EPA labeling categories, or whether the tests had good predictive ability only for a portion of the irritation scale. We found that the latter case was true for the data we analyzed. This led us to develop a testing strategy which utilizes the Cytosensor assay and the EpiOcular assay to identify the mild products, *e.g.,* Categories III and IV, depending on the physical state of the material. Substances which scored more irritating than a Category III were moved to the more robust BCOP assay to determine if the materials were either Category I or Category II materials.

This testing scheme also can begin with the BCOP assay for materials expected from their composition to be highly irritating. However, if the BCOP assay shows the substance to be of a lower (Category III) irritation potential, the substance may be retested in the Cytosensor or EpiOcular assay to determine if it is a Category III or Category IV material. This strategy is depicted in Figure 1-1.

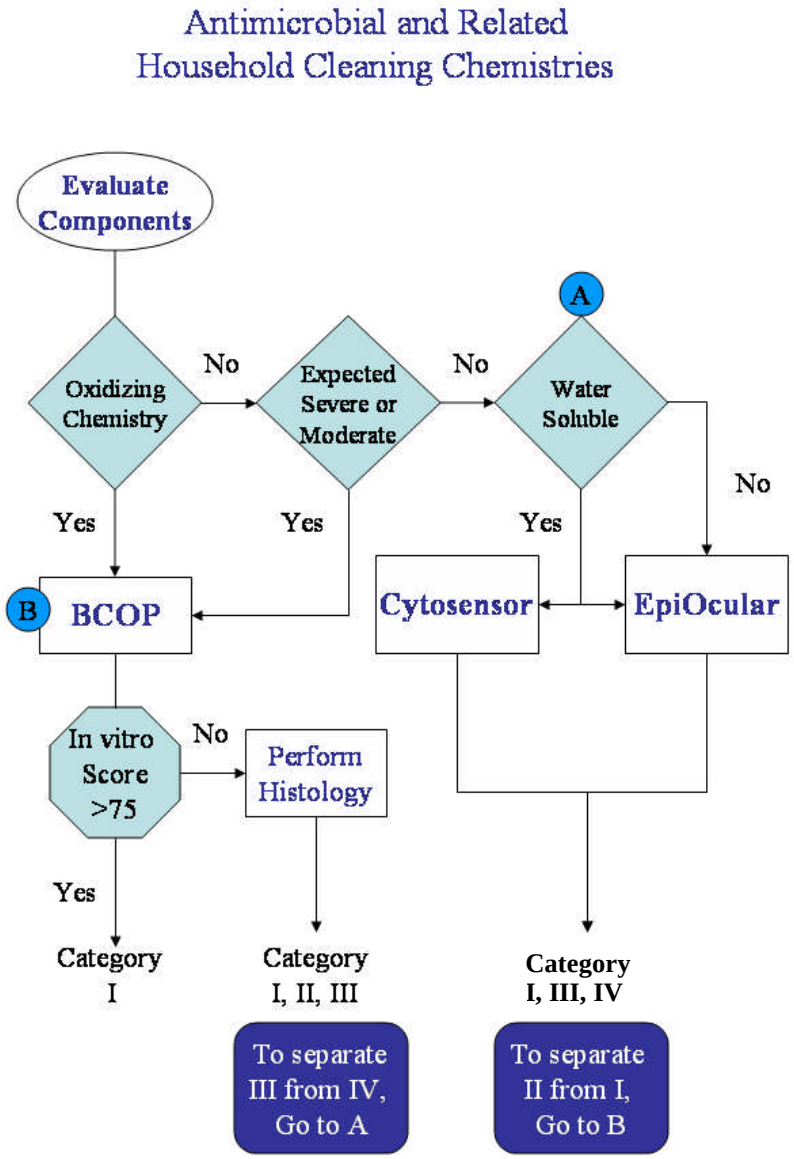


Figure 1-1 The proposed testing strategy for determining the EPA toxicity category for anti-microbial cleaning products.

We wish to make it clear that the above strategy is self-correcting if the initial estimate of irritation potential of a test substance is incorrect. If a highly irritating material is tested in the Cytosensor or EpiOcular assays, it will receive a score indicating that it is a highly irritating (category I) material. If further resolution is desired (to determine if it is actually a Category II material rather than a Category I material), the formulation can then be further tested in the BCOP assay. Similarly a mild material will be identified as a Category III material by the BCOP assay. If it is important to the company to distinguish between a Category III and IV for labeling and marketing

purposes, then an additional Cytosensor or EpiOcular assay may be required to make that determination.

1.1.3 Confidential information

Manufacturers who are participating in this program by submitting data have agreed that any information that is contained in this submission is non-confidential. However, the submitters do desire that individual data not be linked to a specific company. Therefore, that information is not included, and the data are grouped so that no linkage can be made to the company that generated it.

1.2 Regulatory rationale and applicability

1.2.1 Current regulatory testing requirements for which the proposed test method is applicable

The proposed test methods will be used to make labeling decisions for anti-microbial cleaning products as required by the EPA's Office of Pesticide Programs (EPA 2003).

The traditional method of making the labeling decisions is based on the Draize rabbit eye irritation test (Draize, Woodard et al. 1944). In this test, a scoring scheme is applied to the eyes of albino rabbits whose eyes have been exposed to a test material by application within the conjunctival sac. The degree of irritation is classified according to the ocular irritation criteria of Kay and Calandra (1962). This process is described in Acute Eye Irritation (EPA 1998) published in August 1998. The same scoring system is also used for grading and interpretation of data using the Low Volume Eye Test (LVET) method.

As stated in the BRD produced by NICEATM for the BCOP assay: "The EPA ocular irritation classification regulation and testing guidelines (EPA 1998; EPA 2003) are based on the most severe response in one animal in a group of three or more animals. This classification system takes into consideration the kinds of ocular effects produced, as well as the reversibility and the severity of the effects. The EPA classifies substances into four ocular irritant categories, ranging from I to IV (Table 1-1) (EPA 2003). Category I substances are defined as corrosive or severe irritants, while classification from II to IV is based on decreasing irritation severity, as well as the time required for irritation to clear. Irritation that clears in 8 to 21 days is classified as Category II, while irritation that clears within seven days is classified as Category III. For Category IV substances, irritation clears within 24 hours."

Table 1-1 *In vivo* Ocular Irritancy Classification Systems

Regulatory Agency (Authorizing Act)	Number of Animals	Minimum Observations Times (after treatment)	Mean Score Taken?	Positive Response	Irritant/Nonirritant Classification
EPA (FIFRA; TSCA; and The Federal Environmental Pesticide Control Act)	At least 3*	1 hour, 1, 2, 3, 7, 14, and 21 days	No	-Maximum score in an animal used for classification -Opacity or Iritis ≥ 1 or Redness or chemosis ≥ 2	One or more positive animals needed for classification in categories below: I = Corrosive, corneal involvement, or irritation persisting more than 21 days II = Corneal involvement or irritation clearing in 8-21 days III = Corneal involvement or irritation clearing in 7 days or less IV = Minimal effects clearing in less than 24 hours
GHS – Irreversible Eye Effects	3	1, 2, 3, days (observation until Day 21)	Yes	Mean animal values (over Days 1, 2, and 3) of: • Opacity ≥ 3 and/or • Iritis ≥ 1.5	1 = At least 2 positive response animals 1 = At least 1 animal where Opacity, Chemosis, Redness, or Iritis > 0 on Day 21
GHS – Reversible Eye Effects	3	1, 2, 3 days (observation until Day 21)	Yes	Mean animal values (over Days 1, 2, and 3) of: • Opacity or Iritis ≥ 1 or • Redness or Chemosis ≥ 2	2A = At least 2 positive response animals and the effect fully reverses in 21 days 2B = At least 2 positive response animals and effect fully reverses in 7 days

Abbreviations: EPA = U.S. Environmental Protection Agency; FIFRA = Federal Insecticide, Fungicide, and Rodenticide Act; GHS = United Nations Globally Harmonized System; TSCA = Toxic Substances Control Act.

* - Only one animal is required if the result in that animal is corrosive.

Table 1-1 Cont'd

Regulatory Agency (Authorizing Act)	Number of Animals	Minimum Observations Times (after treatment)	Mean Score Taken?	Positive Response	Irritant/Nonirritant Classification
European Union (EU)	**Current Directive:** 1 if severe effects are suspected or 3 if no severe effects are suspected **Prior Directive:** 3 or 6 animals	**1, 2, 3 days (observation until Day 21)**	**Yes**	**<u>6 Animals</u>** **Mean study values (scores averaged over all animals in study over Days 1, 2, and 3) of:** • **Opacity or Chemosis ≥ 2** • **Redness ≥ 2.5** **or** • **Iritis ≥ 1** **<u>3 Animals</u>** **Individual animal mean values (scores for each endpoint are averaged for each animal over Days 1, 2, and 3) of:** • **Opacity or Chemosis ≥ 2** • **Redness ≥ 2.5** **or** • **Iritis ≥ 1**	**R36 =** • **Mean study value (when more than 3 animals are tested) where:** o **2 ≤ Opacity < 3 or** o **1 ≤ Iritis < 1.5 or** o **Redness ≥ 2.5 or** o **Chemosis ≥ 2** • **If 2 of 3 tested animal have individual animal mean values that falls into one of the following categories:** o **2 ≤ Opacity < 3 or** o **1 ≤ Iritis < 1.5 or** o **Redness ≥ 2.5 or** o **Chemosis ≥ 2** **R41 =** • **Mean study value (when more than three animals are tested) where:** o **Opacity ≥ 3 or Iritis > 1.5** • **If 2 of 3 tested animals have individual animal mean values that fall into one of the following categories:** o **Opacity ≥ 3 or Iritis = 2** • **At least one animal where ocular lesions are still present at the end of the observation period, typically Day 21**

1.2.2 Intended regulatory use (*i.e.,* replacement) of the proposed method

The proposed testing scheme is designed to replace the Draize rabbit eye irritation test for the purpose of toxicity labeling of anti-microbial cleaning products (see above).

1.2.3 Similarities between data obtained using this method and the current *in vivo* data

The current *in vivo* data consist of information about the cornea (area and amount of opacity), the iris (iritis) and the conjunctiva (redness and chemosis).

Data obtained from the proposed *in vitro* testing scheme give information about toxicity mainly to the cornea and the conjunctiva. Two of the *in vitro* ocular irritation tests proposed (EO and CM) give information about the direct toxicity of the test material to cells. This is the same type of toxicity that occurs in the outer surface of the cornea and to the conjunctiva. The third *in vitro* ocular test utilizes an excised bovine cornea, and thus the type of initial damage that is seen in this *in vitro* (or *ex vivo*) test is very similar to what occurs to the animal cornea during a traditional eye irritation test.

1.2.4 Fit of method into the overall strategy of toxicity or safety assessment

The proposed *in vitro* testing strategy provides a complete tiered assessment process to determine the EPA toxicity category and product labeling for eye irritation caused by anti-microbial cleaning products.

1.3 Scientific basis for the proposed test method

1.3.1 Purpose and mechanistic basis of the proposed test methods

Data from three *in vitro* methods are used in this submission. These *in vitro* methods – the Cytosensor assay, the EpiOcular assay and/or the BCOP assay – were primarily chosen because they had been extensively used by participating companies to assess cleaning products and similar materials, and because there were *in vivo* data available which could be paired with the *in vitro* data. The mechanistic basis of each of these assays is described in detail below.

1.3.1.1 Cytosensor Microphysiometer (CM) Assay

The Cytosensor is a machine which measures the metabolic activity of a small population of cells grown as a monolayer in a Transwell cup. The cells are exposed to increasing concentrations of a test substance, and their metabolic activity (an estimate of their viability) is measured after each exposure. As the toxicity of the test substance increases, the metabolic activity decreases until

eventually the cells may be completely killed. The endpoint of the assay is the MRD_{50} (concentration of test material which reduces the metabolic rate to 50% of the control rate). The more irritating the test material, the lower the MRD_{50}.

1.3.1.1.1 Intended uses / purpose of the CM

Currently the CM is used by industry early in the new product development process to screen primarily liquid ingredients for cosmetic, personal care, and household cleaning products. This screening is then often followed by evaluations of the final formulations for final in-house safety and labeling decisions. Data from the CM may be combined with information from other *in vitro,* existing *in vivo*, and *in silico* assays on the formulation and/or the ingredients contained within to provide a "weight of evidence" evaluation of the formulation. Information from this assay is generally not combined with new animal data in making the final safety decision for the product.

At the time the CM technology was developed, a number of *in vitro* assays such as the Neutral Red Uptake assay were already proposed as potential replacements for the Draize eye irritation test. However, the great advantage of the CM, or its predecessor the silicon microphysiometer (SM), technology was that measurement could be made of the cytotoxic response of the target cells in real time, as opposed to the 2-3 days or longer time which was required of the existing cytotoxicity assays. Thus, the assay was mainly created not to reveal a completely new endpoint, but rather to provide data in a much shorter time period. Subsequently, it was realized that greater sensitivity of the CM method made it useful in identifying differences between formulations which were already determined to be very mild.

1.3.1.1.2 Regulatory rationale and applicability of the CM

To the best of our knowledge, the CM assay is not currently included in the regulatory scheme of any country. Data are used primarily to evaluate raw materials and formulations where regulatory registration is not required. It has been reviewed informally by regulatory agencies in the US as part of the Interagency Regulatory Alternatives Group (IRAG) evaluation of alternative ocular irritation assays (Botham, Osborne et al. 1997). A BRD on the performance of the CM test method is currently being prepared for review by ECVAM as part of their ocular toxicity method validation program.

1.3.1.1.3 Scientific basis for the CM test

Topical applications of chemicals can kill cells in several ways; among these are lysis of membranes, denaturation of proteins, saponification of lipids, and alkylation or other covalent interactions with macromolecules. The first three modes of action kill or damage very rapidly while the last may act rapidly but the evidence

of the action may take some time to be manifested (Maurer, Parker et al. 2002). Certain chemical classes are associated with these modes of action. Surfactants are primarily associated with membrane lysis although cationic surfactants may also act to precipitate proteins and other macromolecules. Organic solvents can act to delipidize and thus lyse membranes as well as denature (coagulate or precipitate) proteins. Acids tend to coagulate or precipitate proteins. Alkalis saponify lipids and denature proteins in a way that tends to allow them to penetrate into the cornea. Bleaches, peroxides, alkylators (*e.g.,* mustards) bind to macromolecules (especially DNA) leading to cell death.

Damage to the eye is a function of the inherent cytotoxicity potential of the chemical or mixture, the effective concentration impacting the tissues and the residence time at that concentration on or in the tissues. The effective exposure is a combination of concentration and time of exposure (Figure 1-2). For example, a neat organic solvent may have a high cytotoxic potential but if it rapidly evaporates, the effective residence time will be less. Putting a large volume into a closed sac (*e.g.,* lower conjunctival sac of the rabbit eye) will produce a very different effective exposure than a smaller amount placed (or accidentally splashed) onto the open surface of the cornea. Another solvent may have a longer residence time but have its cytotoxic potential rapidly reduced by dilution with tears. In this case, the irritation potential in a species with a low propensity to tear could show much more irritation than in a species with a high propensity to tear. The effective exposure to solids (powders) in the eye is a particular challenge. Powders placed into the conjunctival sac may have a residence time that ranges from minutes to a full day (and longer in some older studies) (Prinsen 2006). Traditional studies of eye irritation potential do not measure or control the effective exposure within or among studies. Thus, efforts to model exposure in alternative test systems are based on best estimates and approximations.

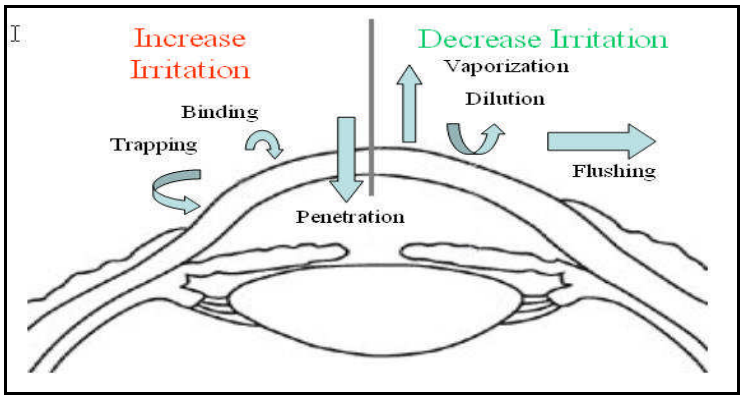


Figure 1-2 Factors that impact exposure to the eye

Mechanistically, this cytotoxicity assay is intended to model the action of the surfactant on the cell membranes of the corneal and conjunctival epithelium where the test article would reside in an *in vivo* exposure. The potency of the surfactant (or surfactant formulation) *in vivo* is related to the area and number of cell layers that can be lysed during the effective exposure period. More potent (and/or more

substantive) surfactants will be more effective at a given concentration and exposure period. Potency can be a function of concentration (*e.g.,* in a formulation) or chemical structure. Thus, a lower concentration of a more potent surfactant or more concentrated formulation would be required to lyse the membranes, and thus kill a given fraction of the cells in the epithelia (both corneal and conjunctival). Expressed another way, a given concentration of a more potent test material should lyse more cells (*i.e.,* greater depth of penetration and injury). Initial depth of injury has been shown by Maurer, Jester, and collaborators (Jester, Petroll et al. 1998; Jester, Li et al. 2001; Maurer, Parker et al. 2002) to relate directly to the degree and duration of ocular injury (Figure 1-3). Their work has shown the relationship between cell initial killing and the resulting irritation. In the cytotoxicity assays with monolayer cells, a similar relationship between potency and effective concentration is expected for killing 50% of the target cell population (Harbell, Koontz et al. 1997).

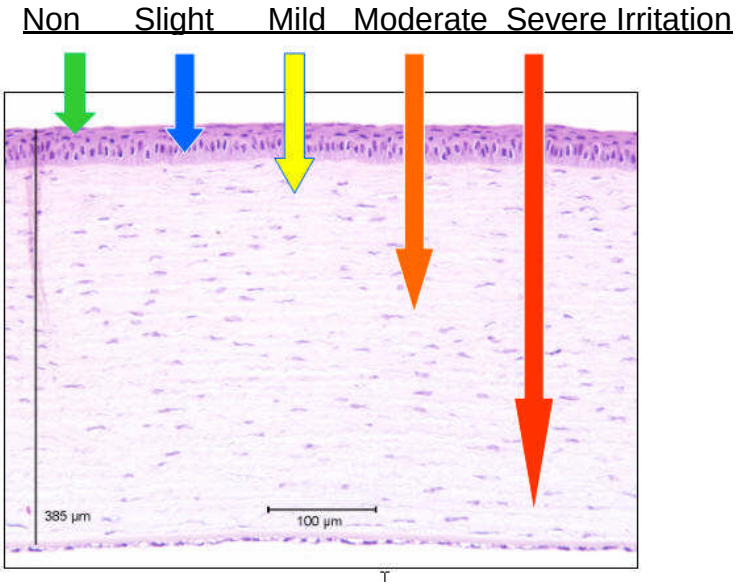


Figure 1-3 Summary of the Depth of Injury Model

The CM estimates the metabolic rate (glucose utilization rate) of a population of cells by measuring the rate of excretion of acid by-products and resulting decrease in pH of the surrounding medium in an enclosed chamber. The rate of change in pH per unit time becomes the metabolic rate of the population. The basal metabolic rate and the ratio of glycolytic to aerobic metabolism (Krebs Cycle) may be different for different cell types. However, for the population of any one cell type, the ratio remains similar if the cells are handled in a consistent fashion. If a test material causes cytotoxicity to this population of cells it is assumed that the metabolic rate will fall. However, the metabolic rate may not fall immediately after exposure of the cells to a dilute concentration of toxicant. Populations of cells in culture are reported to metabolize glucose at only a fraction of their maximal metabolic rate (McConnell, Owicki et al. 1992). Thus, an up regulation of glucose metabolism can occur if the cells need energy to maintain their integrity in the face of a mild biochemical insult. For example, exposure to a subcytotoxic concentration of surfactant can increase membrane leakage (to ions and water). This in turn can

lead to an increase in the activity of ATP-dependent ion pumps and increased glucose metabolism. Thus early points in a killing curve can show increases in metabolic rate of 2- to 3-fold, but this metabolic rate then soon falls below 100% as higher concentrations of test material overwhelm the homeostatic controls within the cells (Figure 1-4).

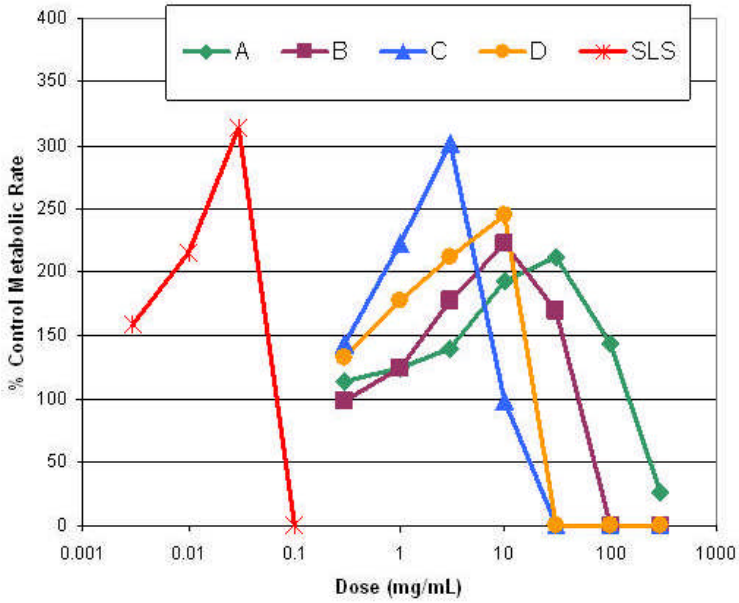


Figure 1-4 Example of the metabolic rate data as a function of surfactant type and concentration

Although the metabolic rate is the physical parameter which is measured during the CM assay, the magnitude of metabolic rate itself is not directly related to eye irritation potential. Rather, the reduction of the metabolic rate to 50% of its basal rate is the parameter used to measure the impact of the test article on the test system (L929 cells in almost all cases). The CM assay exposes a population of cells to increasing concentrations of the test article (diluted in medium). The exposure follows a three step process where the first step is the exposure to the diluted test article, the second is the test article rinse-out and the third is the measurement of the metabolic activity. This means that the impact of the exposure is measured immediately and then a subsequent exposure is performed until the highest testable concentration has been used or the population of cells is severely damaged and the metabolic rate has declined to effectively zero. From the concentration response curve, the concentration that leads to a 50% decline in the metabolic rate of the population (the MRD_{50}) is calculated from the curve. The MRD_{50} values are used to compare test materials and provide a measure of ocular irritancy potential. By current convention, the units of the MRD_{50} are mg/mL.

For ease in understanding the mechanistic basis of the CM assay, a table (Table 1-2) has been compiled describing the events that are commonly considered to occur during eye irritation. Those events that are modeled (or are closely related) by the CM assay are indicated by a Y (yes) indication. It can be seen that the CM

assay most closely models some of the initial stages of interaction of an eye irritant with the cornea. The more distal occurrences in eye irritation such as gross tissue changes in the corneal stroma, and the recovery from the lesions, are not directly modeled. However, if the hypothesis of Jester, Mauer, and others that initial area and depth of injury is predictive of time to, and extent of, recovery, then the measurements made by the CM may have a relationship to recovery as well.

Table 1-2 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses.

Events involved in chemical-induced eye irritation	Modeled by the CM assay?
Chemical interaction with tear film (Klyce and Beuerman 1988; Hackett and McDonald 1994)	N
Chemical binding to the conjunctival epithelium (Hogan and Zimmerman 1962; Hackett and McDonald 1994)	Y
Adhesion molecules compromised (Farquhar and Palade 1963; Van Meer, van Hof et al. 1992; Katahira, Sugiyama et al. 1997)	N
Corneal epithelium damage (Dua, Gomes et al. 1994)	Y
• Inhibition of receptor-mediated membrane transport (Dearman, Cumberbatch et al. 2003)	Y
• Compromise of cell membrane integrity of upper corneal epithelium (Dua, Gomes et al. 1994; Hackett and McDonald 1994; Maurer and Parker 1996)	Y
• Cell membrane lysis of all corneal epithelium layers (Hackett and McDonald 1994)	Y
Hydration of corneal stroma (Hackett and McDonald 1994)	N
Cross-linking of proteins in corneal stroma (Butler and Hammond 1980; Eurell, Sinn et al. 1991; Chan and Hayes 1994)	N
Erosion of corneal stroma (Baldwin, McDonald et al. 1973; Hackett and McDonald 1994; Maurer and Parker 1996)	N
Cell damage to corneal epithelium and limbus (Jacobs and Martens 1990; Wilhelmus 2001)	Partially
Dilation and increased lymphatic leakage from scleral vasculature (Hackett and McDonald 1994)	N
Stimulation of nerve endings, i.e., enhanced blinking, tearing (Chan and Hayes 1994)	N
Erosion of nerve endings in cornea and sclera (Butler and Hammond 1980; Klyce and Beuerman 1988; Araki, Ohahsi et al. 1994)	N
Duration of response, *i.e.,* length of time cell responses deteriorate. Duration of response covers the effects of reactive chemicals which can cause coagulation, saponification, that are effects which develop and increase over time. (Hubert 1992; Maurer and Parker 1996)	N
Recovery from response, *i.e.,* length of time for cell responses to return to control levels (Hubert 1992)	N

1.3.1.2 EpiOcular

The *in vitro* method using the EpiOcular tissue model was developed as a replacement for the Draize eye irritation test (Draize, Woodard et al. 1944; Draize, Woodward et al. 1944). The Draize scoring system is heavily weighted towards corneal damage (80 out of a total of 110 total points) because irreversible damage to the cornea can lead to blindness. Since damage to the cornea is so important both in the Draize scoring scale and to human health, the cornea (specifically its outer surface, the epithelium) is the tissue that is modeled by the EpiOcular tissue model. The EpiOcular protocol models very closely the Low Volume Eye Test (LVET) (Griffith, Nixon et al. 1980) where test materials are applied directly to the surface of the cornea.

The topical application method described in this BRD utilizes a commercially available three-dimensional tissue construct called EpiOcular (Model OCL-200, MatTek Corporation, Ashland, MA) (herein referred to as the EpiOcular tissue model). The EpiOcular tissue model consists of normal, human-derived epidermal keratinocytes that have been cultured to form a stratified, squamous epithelium similar to that found in the human cornea (Figure 1-5). In this model, keratinocytes progressively flatten as the apical surface of the tissue is approached and differentiate to form a multi-layered structure that closely resembles the corneal epithelium *in vivo*. *In vivo*-like growth characteristics are reproduced and include mitotically and metabolically active cells that produce pro-inflammatory growth factors and cytokines important in ocular irritation and inflammation (Thakur, Clegg et al. 1997). Test materials can be applied directly to the surface of the tissue construct to approximate exposure conditions *in vivo*. Damage to the tissue, as reflected by cell cytotoxicity, can be quantified via the chemical reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) and related to a test material's potential for ocular irritation. The current submission describes the relationship between *in vitro* cytotoxicity (time-to-toxicity) and *in vivo* ocular irritation.

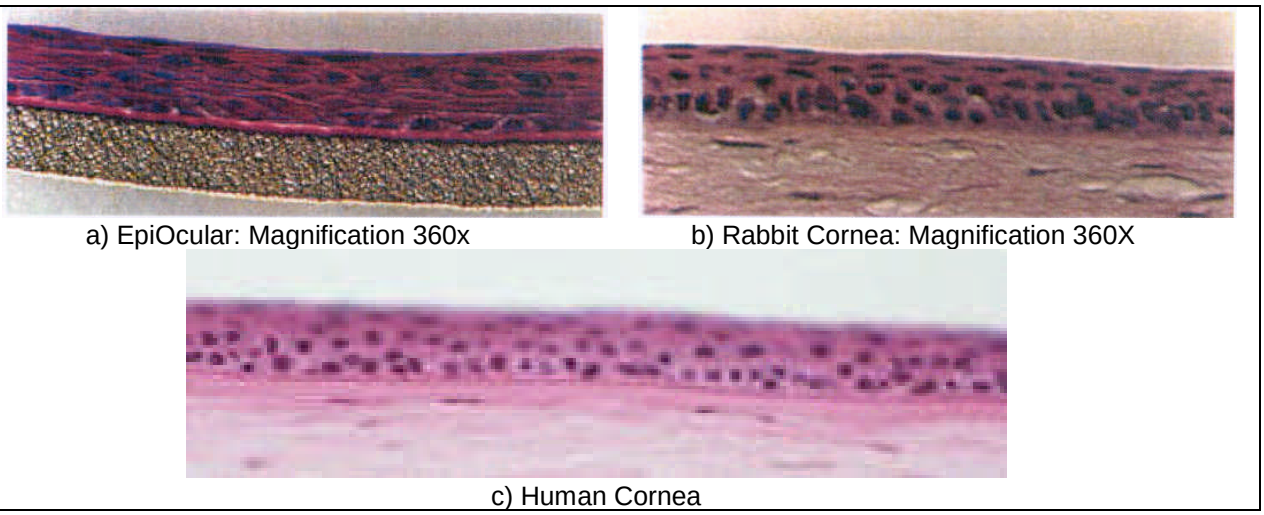


Figure 1-5 Photomicrographs of a) the EpiOcular model showing the stratification and lack of surface keratinization (photo from MatTek Corporation, Ashland, MA), b) the cornea of a rabbit eye (photo courtesy of MatTek Corporation, Ashland, MA), and c) a human cornea.

1.3.1.2.1 Intended uses / purpose of the EpiOcular assay

Very similar to what was described earlier for the CM (Section 1.3.1.1.1), the EpiOcular assay is used by industry early in the new product development process to screen solid or liquid ingredients for cosmetic, personal care, and household cleaning products, as well as assessment of irritation potential of final formulations. One advantage that this method has in comparison to the CM test method is that common product formulations like gels, pastes, creams, and powders are completely compatible with the EpiOcular tissue. Toxicity screening activity is then often followed by further EpiOcular evaluations of the final formulations for final in-house safety decisions. Data from the EpiOcular assay may be combined with information from other *in vitro* or *in silico* assays to provide a “weight of evidence” evaluation of the formulation. Information from this assay is generally not combined with new animal data in making the final safety decision for the product.

1.3.1.2.2 Regulatory rationale and applicability of the EpiOcular test method

To the best of our knowledge, the EpiOcular test method is not currently included in the regulatory scheme of any country. Data are used primarily to evaluate raw materials and formulations where regulatory registration is not required. It is in the process of being reviewed by ECVAM as part of their ocular toxicity method validation program.

1.3.1.2.3 Scientific basis for the EpiOcular test method

As described above, the EpiOcular test method is an attempt to model early changes that occur in the cornea after exposure to a potential eye irritant. The model, as shown in Figure 1-5, closely resembles the non-keratinized squamous epithelium of the mammalian cornea. Because this model is maintained at the air:medium interface, the apical surface is accessible for direct application of test material as might occur during a traditional Draize or LVET rabbit eye test or an accidental human exposure.

Since the damage induced by eye irritants is generally progressive from the corneal epithelium through the stroma and potentially to the endothelium, the EpiOcular assay is able to provide information on the first stages of this progression. As an irritant kills cells as it moves through the corneal epithelium, the cytotoxic progress can be estimated by measuring the loss of MTT reducing activity in the EpiOcular tissue using standardized methods. Although the model only represents the corneal epithelium, (very mild responses would also be reflective of some conjunctival irritation), it can be used to estimate deeper damage into the stroma because of the time-to–toxicity measurements (ET_{50}'s) that are made. The quicker a material kills 50% of the cells in the model the more likely it is to progress to deeper layers of the cornea.

It should be clear from this discussion that the EpiOcular assay is most valuable in addressing the milder end of the irritation scale. Very mild materials may take up to 4 hours to kill 50% of the cells. Thus it is relatively easy to differentiate between the degrees of mildness of two closely related mild substances. However, if extremely irritating materials are used with the EpiOcular assay, the rather thin layer of cells comprising the model is killed quite rapidly (on the order of seconds for extremely toxic materials). When materials act this rapidly, it is extremely difficult to differentiate one very toxic material from another which is only slightly less toxic. Thus the EpiOcular assay has been used most successfully with materials which exist in the lower range of irritancy potential. That fact is borne out by the data in this BRD which show that the EpiOcular assay can be used to identify and differentiate EPA Category III from Category IV materials, while the BCOP assay cannot. Conversely the EpiOcular assay does not seem to be able to differentiate EPA Category II materials from EPA Category I materials as easily as the BCOP assay.

For ease in understanding the mechanistic basis of the EO assay, a table (Table 1-3) has been compiled describing the events that are commonly considered to occur during eye irritation. Those events that are modeled (or are closely related) by the EpiOcular assay are indicated by a Y (yes) indication.

It can be seen that the EpiOcular assay most closely models some of the initial stages of interaction of an eye irritant with the cornea. The more distal occurrences in eye irritation such as gross tissue changes in the corneal stroma, and the recovery from the lesions, are not directly modeled. However, if the

hypothesis of Jester, Mauer, and others that initial area and depth of injury is predictive of time to, and extent of, recovery, then the measurements made by the EpiOcular assay may have a relationship to recovery as well.

Table 1-3 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses.

Events involved in chemical-induced eye irritation	Modeled by the EpiOcular assay?
Chemical interaction with tear film (Klyce and Beuerman 1988; Hackett and McDonald 1994)	N
Chemical binding to the conjunctival epithelium (Hogan and Zimmerman 1962; Hackett and McDonald 1994)	Y
Adhesion molecules compromised (Farquhar and Palade 1963; Van Meer, van Hof et al. 1992; Katahira, Sugiyama et al. 1997)	Y
Corneal epithelium damage (Dua, Gomes et al. 1994)	Y
• Inhibition of receptor-mediated membrane transport (Dearman, Cumberbatch et al. 2003)	Y
• Compromise of cell membrane integrity of upper corneal epithelium (Dua, Gomes et al. 1994; Hackett and McDonald 1994; Maurer and Parker 1996)	Y
• Cell membrane lysis of all corneal epithelium layers (Hackett and McDonald 1994)	Y
Hydration of corneal stroma (Hackett and McDonald 1994)	N
Cross-linking of proteins in corneal stroma (Butler and Hammond 1980; Eurell, Sinn et al. 1991; Chan and Hayes 1994)	N
Erosion of corneal stroma (Baldwin, McDonald et al. 1973; Hackett and McDonald 1994; Maurer and Parker 1996)	N
Cell damage to corneal epithelium and limbus (Jacobs and Martens 1990; Wilhelmus 2001)	Partially
Dilation and increased lymphatic leakage from scleral vasculature (Hackett and McDonald 1994)	N
Stimulation of nerve endings, i.e., enhanced blinking, tearing (Chan and Hayes 1994)	N
Erosion of nerve endings in cornea and sclera (Butler and Hammond 1980; Klyce and Beuerman 1988; Araki, Ohahsi et al. 1994)	N
Duration of response, *i.e.,* length of time cell responses deteriorate. Duration of response covers the effects of reactive chemicals which can cause coagulation, saponification, that are effects which develop and increase over time. (Hubert 1992; Maurer and Parker 1996)	N
Recovery from response, *i.e.,* length of time for cell responses to return to control levels (Hubert 1992)	N

1.3.1.3 BCOP

The test system (target tissue) for the BCOP assay is the isolated bovine cornea obtained as a by-product from freshly slaughtered animals (Figure 1-6). The procedures for preparing and handling the test system were developed by Gautheron *et al.* (1992). The assay measures two important components that are predictive of eye irritation; corneal opacity and permeability (Sina 1994). When necessary, the depth and degree of injury may be assessed by histological evaluation.

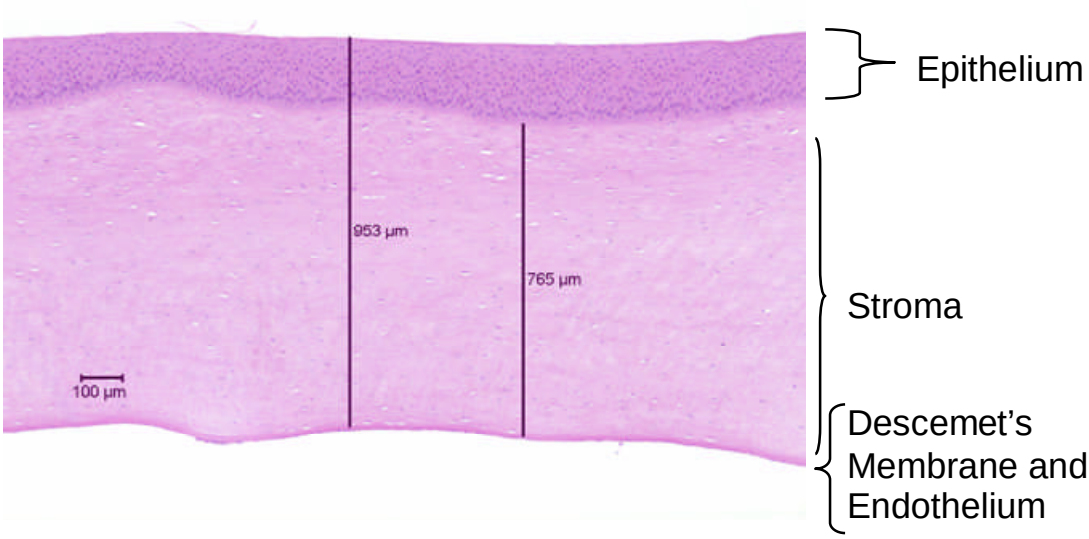


Figure 1-6 A cross-section of a typical bovine cornea as used in the BCOP assay. (H&E stain)

Since the apical surface of the bovine cornea is easily accessible in the organ culture chamber in which the cornea is held, liquid test substances can be easily applied and tested neat unless information about exact in-use (diluted) conditions are desired. Solid test substances are usually tested as a 20% slurry in sterile deionized water. Changes in opacity, permeability to fluorescein, and tissue architecture (depth of injury) are measured and used to assess the relative potential for ocular irritancy of the test substances.

1.3.1.3.1 Intended uses / purpose of the BCOP assay

Very similar to what was described earlier for the CM assay (Section 1.3.1.1.1) and the EpiOcular assay (Section 1.3.1.2.1), the BCOP assay is used by industry early in the product development process to screen solid or liquid ingredients for cosmetic, personal care, and household cleaning products, as well as final formulations. One advantage that this method has in comparison to the CM and EpiOcular test methods is that actual ocular tissue is used in the assay, and, if desired, damage to the cornea can be visualized by conducting histopathological analysis after test article treatment. Often final in-house safety decisions are made based on results from the BCOP assay. Data from the BCOP assay may be combined with information from other *in vitro* or in silico assays to provide a "weight

of evidence" evaluation of the formulation. Information from this assay is generally not combined with new animal data in making the final safety decision for the product.

1.3.1.3.2 Regulatory rationale and applicability of the BCOP test method

To the best of our knowledge, the BCOP test method is not currently included in the regulatory scheme of any country. However, data from the assay that indicates severe irritation has been accepted by regulators from several European Union countries in lieu of animal tests. The test has been reviewed by ICCVAM in their evaluation of the "Current Status of *In vitro* Test Methods for identifying Ocular Corrosives and Severe Irritants." The BRD for the BCOP that was constructed for this effort is appended to this report. We have also quoted freely from this NICEATM report in the preparation of the BCOP portion of this current BRD. The final conclusion of ICCVAM concerning the BCOP assay was that there are sufficient data to support the use of the BCOP test method, in appropriate circumstances and with certain limitations, 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, as part of a weight-of-evidence approach.

Within industry, many toxicologists use results from the BCOP assay (with or without histopathology analysis) to make final safety and labeling decisions for products which do not have formal regulatory registration requirements.

1.3.1.3.3 Scientific basis for the BCOP method

The following discussion of the scientific basis for the BCOP assay is quoted from the NICEATM BRD "Current Status of *In vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method."

"The BCOP is an organotypic model (*i.e.,* isolated whole organ, or component thereof) that provides short-term maintenance of normal physiological and biochemical function of the cornea in an isolated system (Chamberlain, Gad et al. 1997). As noted above, the BCOP was developed as an alternative eye irritation test method in order to obviate the need for laboratory animals as the source for test eyes.

The most commonly used endpoints evaluated in the BCOP assay to measure the extent of damage to the cornea following exposure to a chemical substance are corneal opacity and permeability. Opacity is quantitatively measured by the amount of light transmission through the cornea, and permeability is quantitatively measured as the amount of the small molecule, sodium fluorescein, that penetrates all corneal cell layers. Irritant-induced opacity in the cornea indicates

denaturation/precipitation of proteins in the epithelial or stromal layers and/or swelling, vacuolization, or damage to the cells in the stromal layer (Millichamp 1999). Development of opacity in the cornea, which is normally a transparent tissue, is a significant adverse effect of some irritants that can lead to vision loss. Increased corneal permeability results from damage to the corneal epithelium, which normally serves as a barrier function. In addition, histopathological evaluation of the treated cornea provides useful descriptive information of corneal damage (Curren, Evans et al. 2000; Cooper, Earl et al. 2001).

Histopathology or confocal microscopy would allow for a more accurate assessment of the extent of corneal injury. Maurer *et al.* (2002) proposed that the extent of ocular injury, as measured by confocal microscopy, has the greatest impact on the outcome of such an injury. Live/dead cell staining methods evaluated with confocal microscopy have also been used to determine the extent or depth of corneal injury (Maurer, Li et al. 1997) and in an *ex vivo* corneal button assay (Jester, Li et al. 2001). These studies prompted the authors to suggest that the extent of corneal injury could be used as the basis for developing alternative methods to predict the level of damage produced by ocular irritants." Thus, the BCOP offers the possibility of using depth-of-injury analysis through histopathology to predict the potential outcome of eye injury produced by ocular irritants.

For ease in understanding the mechanistic basis of the BCOP assay, a table (Table 1-4) has been compiled describing the events that are commonly considered to occur during eye irritation. Those events that are modeled (or are closely related) by the BCOP assay are indicated by a Y (yes) indication.

It can be seen that the BCOP assay closely models not only most of the initial stages of interaction of an eye irritant with the cornea, but also some of the more distal occurrences in eye irritation such as gross tissue changes in the corneal stroma. However, the short time period that the cornea can be kept in organ culture limits the amount of recovery, if any, which may occur. Again, if the hypothesis of Jester, Mauer, and others that initial area and depth of injury is predictive of time to, and extent of recovery, then the measurements made by the BCOP assay may have a relationship to recovery as well.

Table 1-4 Summary of events involved in chemical-induced eye irritation *in vivo.* Text in italics represents irreversible responses.

Events involved in chemical-induced eye irritation	Modeled by the BCOP assay?
Chemical interaction with tear film (Klyce and Beuerman 1988; Hackett and McDonald 1994)	N
Chemical binding to the conjunctival epithelium (Hogan and Zimmerman 1962; Hackett and McDonald 1994)	Y
Adhesion molecules compromised (Farquhar and Palade 1963; Van Meer, van Hof et al. 1992; Katahira, Sugiyama et al. 1997)	Y
Corneal epithelium damage (Dua, Gomes et al. 1994)	Y
• Inhibition of receptor-mediated membrane transport (Dearman, Cumberbatch et al. 2003)	Y
• Compromise of cell membrane integrity of upper corneal epithelium (Dua, Gomes et al. 1994; Hackett and McDonald 1994; Maurer and Parker 1996)	Y
• Cell membrane lysis of all corneal epithelium layers (Hackett and McDonald 1994)	Y
Hydration of corneal stroma (Hackett and McDonald 1994)	Y
Cross-linking of proteins in corneal stroma (Butler and Hammond 1980; Eurell, Sinn et al. 1991; Chan and Hayes 1994)	Y
Erosion of corneal stroma (Baldwin, McDonald et al. 1973; Hackett and McDonald 1994; Maurer and Parker 1996)	Y
Cell damage to corneal epithelium and limbus (Jacobs and Martens 1990; Wilhelmus 2001)	Y
Dilation and increased lymphatic leakage from scleral vasculature (Hackett and McDonald 1994)	N
Stimulation of nerve endings, i.e., enhanced blinking, tearing (Chan and Hayes 1994)	N
Erosion of nerve endings in cornea and sclera (Butler and Hammond 1980; Klyce and Beuerman 1988; Araki, Ohahsi et al. 1994)	N
Duration of response, *i.e.,* length of time cell responses deteriorate. Duration of response covers the effects of reactive chemicals which can cause coagulation, saponification, that are effects which develop and increase over time. (Hubert 1992; Maurer and Parker 1996)	Partially
Recovery from response, *i.e.,* length of time for cell responses to return to control levels (Hubert 1992)	N

2 Test Method Components

2.1 Overview of the proposed testing approach

A general review of how this project was structured and how the testing approach was determined has been presented in Section 1 – Introduction and Rationale. The testing approach itself is presented in Figure 1-1 and relies on using one of three *in vitro* assays potentially supplemented with a second *in vitro* assay to further refine the appropriate labeling category.

Anti-microbial cleaning products can be formulated in different ways. Although to begin this study we characterized the formulations into several different classes, i.e. acids, bases, surfactants, solvents, and oxidizing chemistries, we found that most of these classes reacted similarly in the in vitro assays. We eventually concluded that only those materials with oxidizing chemistry and those with a high solvent concentration (>5%) should be treated somewhat differently from the others. It is also useful to determine the water solubility of the formulation since only fully water soluble materials can be tested in the Cytosensor Microphysiometer.

The proposed testing strategy (see Figure 1-1) begins by evaluating the components of the formulation. If the formulation is characterized as having oxidizing chemistry, then the first step is to test it using the BCOP assay. This is done because the oxidizers seem to be overpredicted in the other assay systems (see Section 6 – Test Method Predictive Capacity). Any of the other types of formulations may also be tested in the BCOP assay, although we suggest that formulations thought to be mild or non-irritating (e.g. EPA labeling categories III or IV) be tested first in either the Cytosensor or EpiOcular assays. This is suggested since the latter two assays are better able to identify EPA IV materials than the BCOP assay (see Section 6 – Test Method Predictive Capacity). Conversely, if the formulation is thought to be a strong eye irritant, (e.g. EPA I or II) it is suggested that it first be tested in the BCOP assay. If the formulation is characterized as a high solvent (>5%) product, the BCOP assay should be conducted with a 3 minute exposure rather than the traditional ten minute exposure. This is because our studies showed that some high solvent materials were overclassified by the BCOP if the longer exposure was used (see discussion in Section 6.3.2.2.3).

Table 2-1 describes the BCOP assay *in vitro* score cut-off values for the EPA category designations. If the testing results in a BCOP *in vitro* score that is ≥75 it is given a Category I designation. If testing results in a score ≥25, it is initially given a Category II designation, but histopatholgy of the corneas is conducted to verify the designation (see Section 6.3.3). Similarly, a material scoring <25 (Category III) should have histopathology performed to verify its designation, or it could be retested in the Cytosensor or EpiOcular assays to determine whether it was actually a Category IV rather than a Category III.

Table 2-1 BCOP *in vitro* score and EPA category designation

BCOP *In vitro* Score	EPA Category
in vitro score ≥ 75	Category I
75 > *in vitro* score ≥ 25	Category II (Histopathology should be performed)
in vitro score < 25	Assume Category III (Histopathology should be performed) or retest in Cytosensor or EpiOcular to determine if Category III or IV

When conducting the BCOP assay the following conclusions from Section 6 should be considered:

1) **In general, when testing anti-microbial cleaning product formulations, the BCOP assay should be conducted with a ten minute exposure.**

2) **If the anti-microbial cleaning product contains a solvent at the level of 5% or greater, it should be tested with a three minute exposure.**

3) **All anti-microbial cleaning products having an *In Vitro* Score ≥75 should be classified as an EPA Category I or a GHS Category 1. <u>No histopathology needs to be conducted.</u>**

4) **Anti-microbial cleaning products having an *In Vitro* Score <75 and ≥ 25 are given a preliminary classification of EPA Category II or GHS Category 2A. They should be further assessed with a histopathological evaluation and given the final categorization of whichever determination (*In Vitro* Score or histological evaluation) is more severe.**

5) **Anti-microbial cleaning products having an *In Vitro* Score <25 are given a preliminary classification of EPA Category III or GHS Category 2B. They should be further assessed with a histopathological evaluation and given the final categorization of whichever determination (in vitro score or histological evaluation) is more severe.**

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA III or GHS 2B is actually an EPA IV or a GHS NI, it should be further tested in either the Cytosensor or EpiOcular assays.**

For materials not characterized as having oxidizing chemistry and not suspected to be a severe irritant, either the Cytosensor or EpiOcular test is chosen. Liquids and aqueous soluble materials can be tested with the Cytosensor. Granular, non-aqueous soluble materials and liquid, aqueous soluble materials can be tested in the EpiOcular assay. The choice, other than considering the water solubility requirement of the Cytosensor, would be based solely on the experience of the user with one method or the other. However, as discussed in Section 2.2.1, it is likely that in a few years the Cytosensor assay may no longer be available since its manufacturer is no longer supporting the instrument. At that time the EpiOcular assay (or a similar three-dimensional tissue model) will be the only *in vitro* model available to identify EPA Category IV materials – unless another assay is found in the meantime that can be shown to reliably identify the extremely mild materials.

Using cut-off values for either the Cytosensor or EpiOcular assays that are described later in this submission, a decision can be made whether the material is a Category IV, III, or I. Both of these tests were designed to evaluate mild materials and although both can identify severe materials, they do not have the ability to discriminate between Category I and Category II materials. If there is a desire to differentiate between Category I and II materials the BCOP assay must be used.

When conducting the Cytosensor assay the following conclusions from Section 6 should be considered:

1) **Anti-microbial cleaning products having an oxidizing chemistry should not be tested with the Cytosensor assay.**

2) **Only fully water soluble anti-microbial cleaning products can be tested with the Cytosensor assay.**

3) **If the anti-microbial cleaning product has an MRD_{50} score of <2 mg/ml, it is classified as EPA Category I or GHS Category 1.**

4) **If the anti-microbial cleaning product has an MRD_{50} score of ≥2 mg/ml, but < 80 mg/ml, it is classified as EPA Category III. If the anti-microbial cleaning product has an MRD_{50} score of ≥2 mg/ml, but <10 mg/ml, it is classified as GHS Category 2B.**

5) **If the anti-microbial cleaning product has an MRD_{50} score of ≥80 mg/ml, it is classified as EPA Category IV. If the anti-microbial cleaning product has an MRD_{50} score of ≥10 mg/ml, it is classified GHS Category NI.**

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA I or GHS 1 is actually an EPA II or a GHS 2A, it should be further tested in the BCOP assay.**

When conducting the EpiOcular assay the following conclusions from Section 6 should be considered:

1) **Anti-microbial cleaning products having an oxidizing chemistry should not be tested with the EpiOcular assay.**

2) **Both water soluble and water insoluble anti-microbial cleaning products can be tested with the EpiOcular assay.**

3) **If the anti-microbial cleaning product has an ET_{50} score of <4 minutes, it is classified as EPA Category I or GHS Category 1.**

4) **If the anti-microbial cleaning product has an ET_{50} score of ≥4 minutes, but <70 minutes, it is classified as EPA Category III or GHS Category 2B.**

5) **If the anti-microbial cleaning product has an ET_{50} score of ≥70 minutes, it is classified as EPA Category IV or GHS Category NI.**

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA I or GHS 1 is actually an EPA II or a GHS 2A, it should be further tested in the BCOP assay.**

The above strategy - which provides the option for using several different *in vitro* methods – was devised because we found that no single *in vitro* test was able adequately cover the entire range of irritation that is covered by the EPA labeling categories I – IV. The BCOP is a more robust tissue and is able to differentiate the more aggressive materials from each other, while the CM and EO are more sensitive methods and thus better able to resolve differences between milder materials. **Note that if the original decision that the test material falls in the severe range or in the mild range proves to be false when the material is actually tested, the strategy still works; the testing may just take longer because a second assay may have to be used.** If a mild material is mistakenly put into the BCOP it will be identified as a Category III (remember: the BCOP cannot differentiate a IV from a III, and in such a case the more conservative category must be given). To determine if this material is a IV, a second assay in Cytosensor or EpiOcular would have to be conducted.

Similarly if a severe material is tested in the EpiOcular or Cytosensor assays it will be identified as a Category I. If it is necessary to find out if it's actually a Category II, it must be retested in the BCOP. <u>The strategy is self-correcting so there is no worry about initially choosing an incorrect test method using this approach</u>.

2.2 Detailed description and rationale for each assay

The methodologies utilized in the proposed *in vitro* strategy for toxicity classification of anti-microbial cleaning products are the Cytosensor assay, the EpiOcular assay, and the BCOP assay. The methodology used for each of these assays is described below.

2.2.1 Overview of how the CM test method is conducted

The CM uses a low volume flow-through chamber and a light-addressable potentiometer to measure the metabolic rate of a cell population. Metabolic rate is determined indirectly by the number of protons excreted into the low buffer medium (change in pH) per unit time. The light-addressable potentiometer forms the bottom of the flow-through chamber and serves as a very sensitive and stable pH meter. While medium is flowing through the chamber, the pH is stable and governed by the medium. When the flow of medium is stopped, the pH begins to drop in a linear fashion over time. The actual change in pH during this measurement is generally less than 0.2 pH units.

Data contained in this BRD were generated with two different instruments. One was the predecessor instrument to the current CM, the Silicon Microphysiometer (SM). In the SM (Figure 2-1) target cells were grown on a glass coverslip and the coverslip was inverted over the top of the sensor chip to form a flow-through chamber (Figure 2-2). A minority of data was generated with the SM protocol. The majority of the data in this BRD were generated with the Cytosensor.

The positive control currently used for CM studies at IIVS is SLS (using a stock concentration of 100 mg/mL in water). The current (as of 4/28/08) acceptable MRD_{50} is 79.8 µg/mL± 11.3 µg/mL.

To conduct the Cytosensor protocol as used for the majority of studies reported in this BRD (see Annex A1), cells are grown on a Transwell membrane (discussed below). The whole Transwell is placed into the sensor chamber and a plunger (with a spacer) pressed down on the membrane to seal it. The sensor chamber is composed of the light-addressable potentiometer sensor (sensor chip) on the bottom and ports for the medium (inlet and outlet). There is a small medium-filled space between the sensor chip and the bottom of the Transwell. The cells are attached to the top of the membrane so that the acid metabolites must pass through the membrane pores to reach the space in the lower part of the chamber. The medium is passed over the cells on the upper side of the membrane. Figure 2-3 shows the operating components of the instrument and Figure 2-4 shows the low volume sensor chamber (Transwell configuration). Based on the comparison of data generated in both the SM and CM, Procter & Gamble established a conversion algorithm so that all results generated initially from the SM could be compared to the results generated with the CM (details provided in section 2.2.1.1).

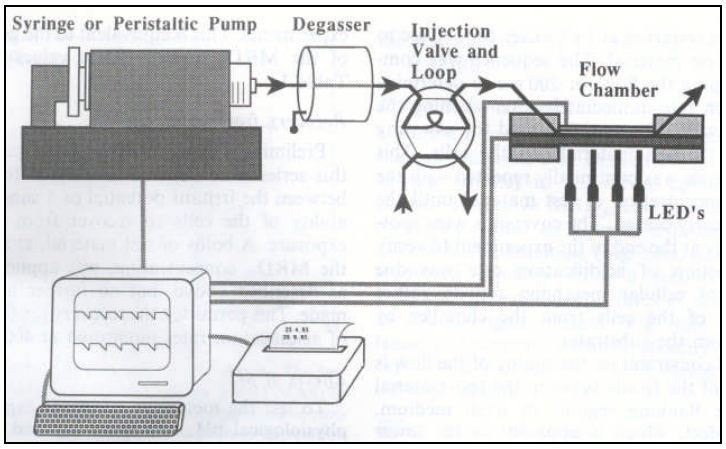


Figure 2-1 Diagram of the operating components of the silicon microphysiometer (Bruner, Miller et al. 1991)

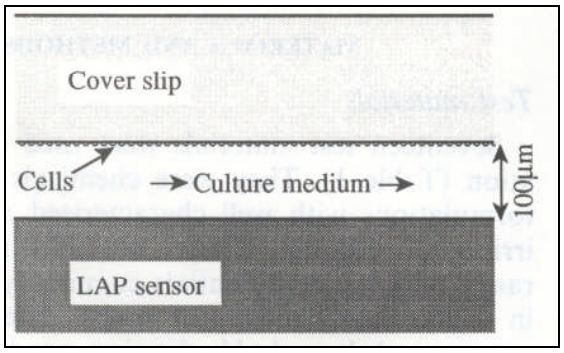


Figure 2-2 The original silicon microphysiometer sensor chamber with the coverslip in place (Bruner, Miller et al. 1991)

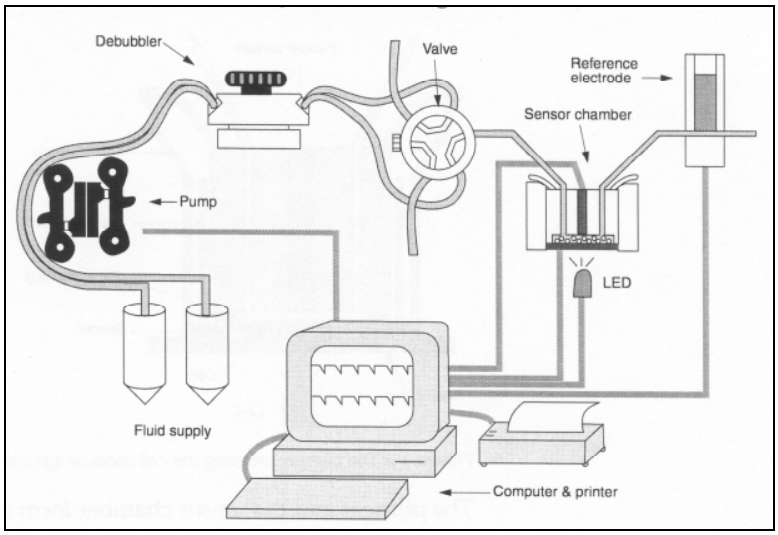


Figure 2-3 Diagram of the operating components of the Cytosensor (Cytosensor Manual)

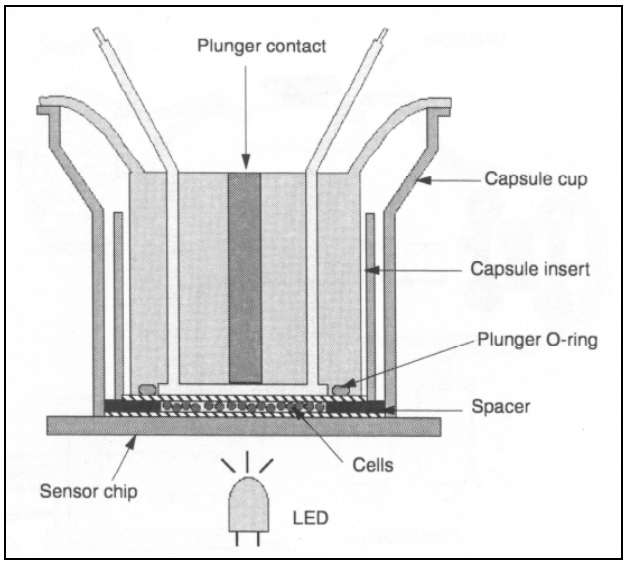


Figure 2-4 The Cytosensor chamber with the Transwell in place (Cytosensor Manual)

Originally, the silicon microphysiometer (coverslip chamber) used a 15-minute exposure, rinse, and read cycle. The cells were exposed to each concentration in two phases. In the first phase, the diluted test article was pumped (1.67 µL/sec) through the chamber for 120 seconds and then the flow halted for 200 seconds (total of 320 seconds of exposure). The chamber was then rinsed with fresh medium at the same rate for 380 seconds. The flow was then stopped for 200 seconds while the acidification rate was measured. This exposure protocol was used primarily on normal human epidermal keratinocytes (Bruner, Miller et al. 1991). Most of the studies in this BRD used L929 cells as the test system. The exposure protocol was altered so that the cells were exposed to the test article for a total of 500 seconds (300 seconds of flow and 200 seconds with the flow off), rinsed for 400 seconds, and the metabolic rate determined for 169 seconds. Flow was restarted with medium before the next dose was introduced. Because the valves were turned manually, the total cycle time was 1100 seconds.

In contrast, the Cytosensor (both the commercial instrument and the silicon microphysiometer with “Cytosensor-like” chambers used a 20-minute (1200-second) exposure, rinse, and read cycle. This is still the current protocol. The cells are exposed 810 seconds (100 µL per minute for one minute and 20 µL per minute for 12.5 minutes). The rinse cycle lasts for 6 minutes and the flow is 100 µL per minute. Finally, the flow is stopped for 25 seconds and the change in pH is measured. For the purposes of the BRD, this will be the standard Transwell protocol (for either the converted silicon microphysiometer or the Cytosensor).

The bulk of the available data come from the Transwell protocol using the 810-second exposure. The Transwell was introduced by Molecular Devices, Inc. to allow more efficient introduction of the test system to the sensor chambers (including non-adherent cells in a gelatin matrix). However, this change limited the cell density and types of cells that could be used. The Transwells have 3 micron pores that allow efficient communication between the upper surface of the

membrane (with the cells) and the lower surface that faces the sensor itself. Confluent cell layers would interfere with this communication and so the cell density was reduced to a standard $6x10^5$ cells per well (seeded the day before use). The Transwell uses a polycarbonate filter membrane that is less prone to interaction with test materials than other types of membranes but does not allow the human keratinocytes to attach. Thus, the L929 cells were selected because they would readily attach and were easy to grow in continuous culture. With the change to L929 cells, the SM exposure protocol was changed to 500 seconds. This is the protocol that was used for most of the SM studies in this BRD. This is also the same protocol that was used in the IIVS positive control database before a switch was made to the CM.

2.2.1.1 Development of Conversion Algorithm between SM and CM

At the time that the SM was replaced with the CM by Molecular Devices, Inc., The Procter & Gamble Company sponsored a study to compare data obtained with the SM (coverslip protocol) for a set of 11 surfactant-containing materials with data obtained for the same materials with the CM (Transwell protocol). The studies were carried out concurrently at a single laboratory (Microbiological Associates, Inc.). The testing protocol utilized a preliminary trial followed by at least three definitive trials. Data produced by the SM and CM are shown in Tables 2-2 & 2-3, respectively. It can be seen that the overall mean CV for each of the two methods is very similar (22.8% for the SM; 21.8% for the CM).

Following data collection from both instruments, the data were compared and the following equation was derived to translate SM coverslip data to CM Transwell data:

$$\text{Log}_{10}\,(\text{Cytosensor MRD}_{50}) = 0.135 + 0.7753 \times \text{Log}_{10}(\text{Silicon Microphysiometer MRD}_{50}).$$

A graph depicting the relationship between the SM and CM is given in Figure 2-5. The current standard Cytosensor protocol is attached in Annex A1.

Table 2-2 Silicon Microphysiometer data for 11 surfactant-containing materials from P&G

Substance	Prelim*	Trial 1	Trial 2	Trial 3	Trial 4	Mean MRD_{50} (mg/mL)	SD	CV (%)
#1	21.368	18.116	25.510	20.408		21.345	3.785	17.7
#2	+	0.083	0.085	0.082		0.083	0.001	1.7
#3	+	0.291	0.266	0.263		0.273	0.015	5.5
#4	+	0.247	0.153	0.435	0.298	0.283	0.117	41.5
#5	+	13.643	13.004	9.434		12.027	2.268	18.9
#6	+	0.042	0.027	0.026		0.032	0.009	28.2
#7	0.161	0.093	0.139	0.198		0.143	0.053	36.8
#8	0.714	2.020	1.239	1.595		1.618	0.391	24.2
#9	0.094	0.043	0.032	0.039		0.038	0.006	14.7
#10	0.020	0.045	0.038	0.026		0.036	0.010	26.9
#11	+	0.081	0.094	0.152		0.109	0.038	34.5
Mean								**22.8**
Median								**24.2**

* Not included in the mean calculation
+ Value not determined during assay

Table 2-3 Cytosensor Microphysiometer data for 11 surfactant-containing materials from P&G

Substance	Prelim*	Trial 1	Trial 2	Trial 3	Trial 4	Mean MRD_{50} (mg/mL)	SD	CV (%)
#1	90.909	56.497	48.544	62.500		55.847	7.001	12.5
#2	0.223	0.254	0.424	0.283		0.320	0.091	28.4
#3	0.758	0.794	0.552	0.820		0.722	0.147	20.4
#4	0.452	0.442	0.412	0.431		0.428	0.016	3.7
#5	19.120	9.091	11.429	5.319		8.613	3.083	35.8
#6	0.067	0.074	0.052	0.075		0.067	0.013	19.2
#7	0.251	0.177	0.288	0.267		0.244	0.059	24.3
#8	2.288	2.110	2.016	2.457		2.194	0.232	10.6
#9	3.497	1.475	4.367	3.802		3.215	1.533	47.7
#10	0.282	+	0.139	0.151	0.165	0.152	0.013	8.5
#11	0.251	0.268	0.159	0.281		0.236	0.067	28.4
Mean								**21.8**
Median								**20.4**

* Not included in the mean calculation
+ Value not determined during assay

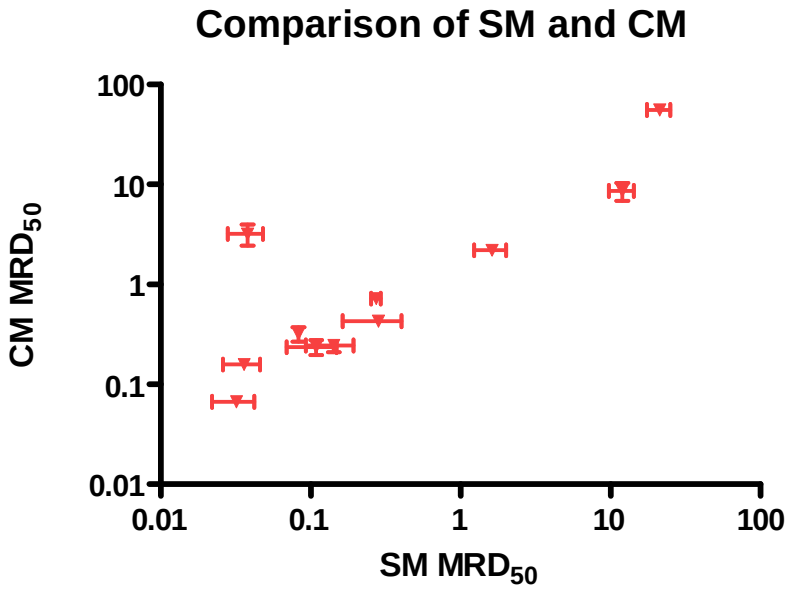


Figure 2-5 A comparison of data obtained from 11 surfactant-containing products with SM and CM.

A more complete description of the Cytosensor is given in a Background Review Document recently prepared under contract to ECVAM. Because this BRD is still in the review process it could not be directly appended to this document, but it is quoted from extensively in this BRD. It will be referred to repeatedly in this submission where more detail is required.

2.2.2 Overview of how the EpiOcular test method is conducted

2.2.2.1 Preparation of the EpiOcular tissue (Description provided by the manufacturer, MatTek Corporation, Ashland, MA)

The EpiOcular model is prepared using proprietary manufacturing techniques in which normal human neonatal foreskin keratinocytes, derived from a single donor, are grown under standardized conditions to produce a highly uniform, reproducible cornea-like tissue. The keratinocytes are expanded in monolayer culture and harvested using trypsinization according to standard techniques described in literature available from Cascade Biologics, Inc. (Portland, OR), the commercial vendor from which the keratinocytes are currently obtained. Single cell suspensions of keratinocytes are aliquoted into 10-mm ID Millicell® PCF cell culture inserts (Millipore Corporation, Bedford, MA); polycarbonate Nunc™ cell culture inserts (Nalge Nunc International, Rochester, NY) also serve as suitable substrates. The inserts are placed in a 37ºC, 5% CO_2 incubator and cultured at the air liquid interface, *i.e.,* only the basal side of the cell culture inserts is exposed

to the medium (see Figure 2-6). The culture medium is Dulbecco's Modified Eagle's Medium (DMEM) to which a proprietary mixture of nutrients, growth factors, and hormones has been added; all media are serum free. After approximately one week of culture, the cell culture inserts containing the stratified tissue are placed atop DMEM-enriched agarose gel in a 24-well tissue culture plate. This 24-well plate is hermetically sealed ("packaged") and shipped for commercial sale or stored at 4°C for 24-72 hours prior to its use for testing. For commercial purposes, these packaged tissues are shipped every Monday on wet ice (c.a. 4ºC) via overnight express delivery.

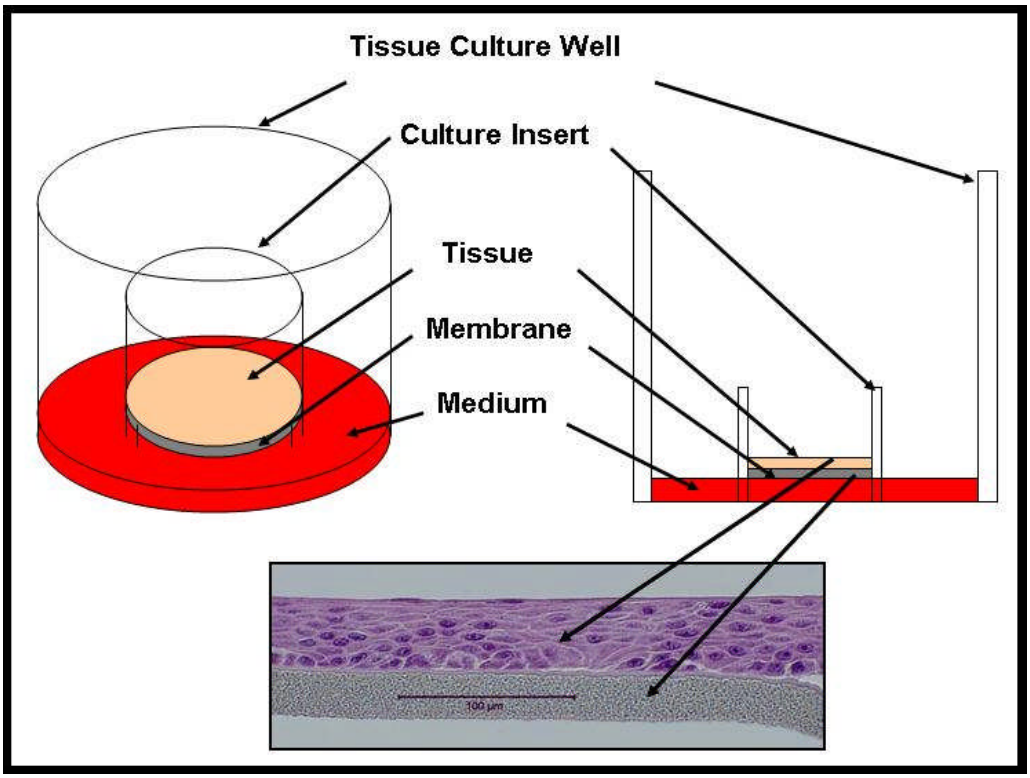


Figure 2-6 Diagrammatic representation of EpiOcular tissue growing in a milliicell chamber placed within a well of a 24-well plate. A photomicrograph of a cross section through the tissue and underlying membrane is included.

2.2.2.2 Test methodology

The protocol used for the majority of EO studies in this BRD can be found in Annex A3. On arrival at the laboratory, EpiOcular tissues are examined for obvious defects and may be rejected based on blistering, excess fluid on the tissue (evidence of an incomplete barrier), air bubbles below the tissue insert, etc. Tissues can be used within 48 hours of receipt. Prior to test article dosing, tissues are transferred (using sterile technique) to 6-well plates that contain fresh assay medium. The tissues are incubated at standard conditions (5% CO_2, 37°C, 95% humidity) for at least 1 hour before use.

EpiOcular tissues which are not used immediately should be equilibrated by placement into a 5% CO_2 environment and stored at 4°C. Experience indicates that repeated equilibration at 5% CO_2, 37°C, 95% humidity (*i.e.,* tissue culture incubator) can produce variability in tissue performance. Prior to dosing with test materials or controls, the tissues are re-fed with fresh, prewarmed assay medium and generally dosed within 30 minutes of refeeding.

The positive control currently used for EO studies at IIVS is 0.3% TRITON® X-100 in water. The current (as of 4/28/08) acceptable ET_{50} is 27.3 min ± 5.0 min..

Dosing of aqueous or semi-viscous test materials is performed with a positive displacement pipette. Solid materials are “sprinkled” onto the surface of the tissue. A dosing device (*e.g.,* the flat end of a sterile push pin) can be used to ensure that the test material covers the complete tissue surface. After application of the test material, the tissues are incubated at standard conditions for various amounts of time estimated to cover the time at which the test material causes 50% toxicity to the tissues. Exposure times generally range from 1 minute to 24 hours. Figure 2-7 presents diagrammatically the procedures used in the EpiOcular assay.

At the end of the incubation period the tissues are removed from the incubator, and the test material is removed from the tissue surface using phosphate buffered saline (PBS). The PBS is sprayed against the Millicell® wall to create a gentle vortex which aids in test material removal. The tissues are then “soaked” in medium at room temperature to ensure a more complete removal of any remaining test material. Following the soak process, the tissues are rinsed again with PBS prior to the MTT reduction step. Complete test material removal is necessary to prevent prolonged exposure and an erroneous estimate of toxicity. Individual tissues are placed into wells containing unreduced *3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide* (MTT) solution. The tissues are incubated at standard conditions for 3 hours. Viable tissue reduces the colorless MTT solution to a dark blue or purple color.

Following exposure to MTT, the tissues are removed and placed into isopropanol for 2 hours at room temperature to extract the reduced MTT. Extracted MTT is thoroughly mixed and transferred to a 96-well plate. The amount of MTT/ethanol in each well is then quantified using a microplate reader. Raw OD_{550} values are used to calculate the final ET_{50} values which are reported in minutes.

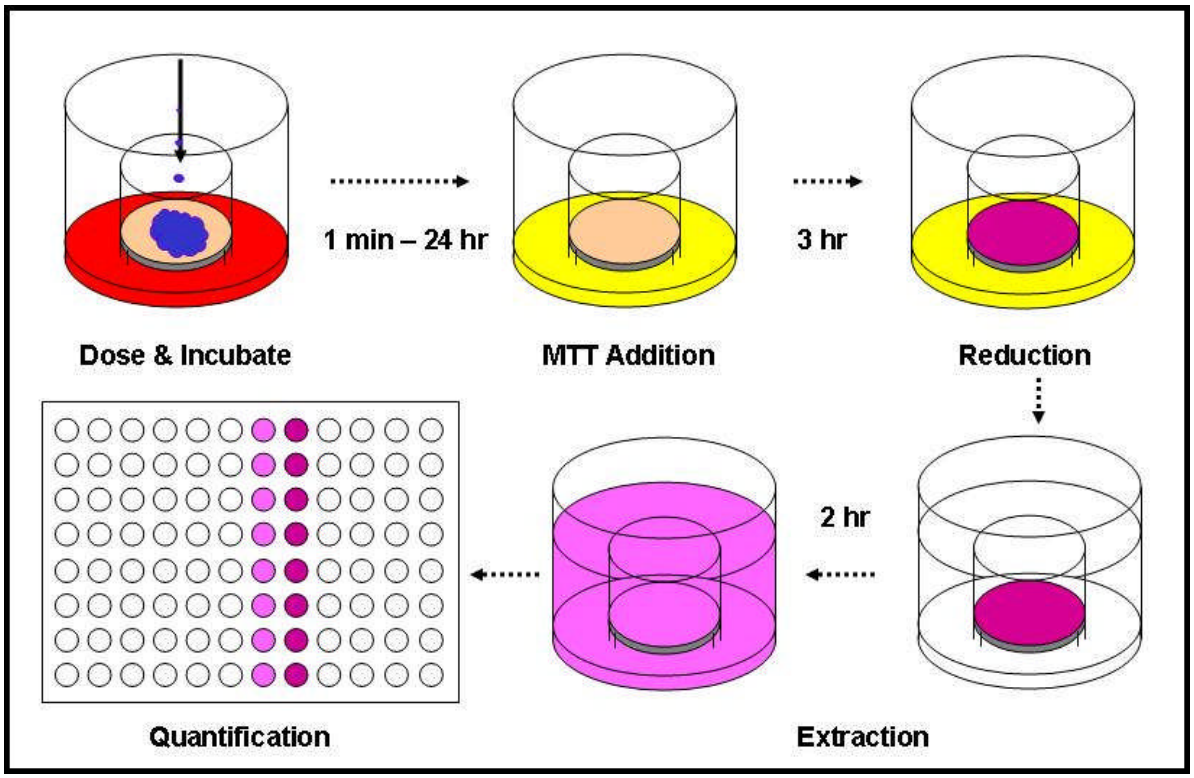


Figure 2-7 Diagrammatic representation of the testing procedure using EpiOcular tissue. Incubation is carried out at 37°C, and test material is thoroughly removed before the addition of MTT.

One technical detail of the assay that can cause serious underestimation of toxicity, and therefore must be carefully controlled, is the possible reduction of MTT by a test material which itself has reducing properties (Liebsch, Traue et al. 2000). If a test material has reducing properties and it binds to the tissue or underlying membrane such that it is not removed during the washing step, then it may reduce the MTT solution resulting in a masking of toxicity to the EpiOcular tissue. This would result in an underprediction of the toxicity category for the test material. This situation can be addressed by screening all test materials for the presence of reducing activity by incubating them directly in MTT solution. If they have reducing properties they will turn the solution purple (see top middle photograph in Figure 2-8). If direct reduction is observed, its actual effect on the assay can be determined by conducting a sham exposure on EpiOcular tissue that has been freeze–killed. If no MTT reduction is seen, then no test material remained on the tissue or membrane after the wash step and the reducing properties of the test material are not of a concern. However, if reduction has occurred the amount can be calculated and that value can be subtracted from the MTT reduction at the identical time point in the full assay so that the true viability of the tissue can be determined.

Other aspects of the assay that can be visualized are shown in Figure 2-8. For example, the photograph in the top left illustrates the results of testing a material with hydroscopic properties. Almost all the medium has been absorbed by the test

material likely causing toxicity to the EpiOcular tissue which might not occur in an *in vivo* situation. Similarly artifactual results can occur unless the presence of air bubbles under the membrane is carefully monitored (Figure 2-8 top right photograph). Large air bubbles can significantly block the passage of MTT into the tissue.

The lower row of photographs in Figure 2-8 demonstrate that the viability of the tissue can be visualized at the conclusion of the MTT exposure step, and therefore these recorded observations of toxicity can be compared to subsequent viability values calculated from the absorbance values.

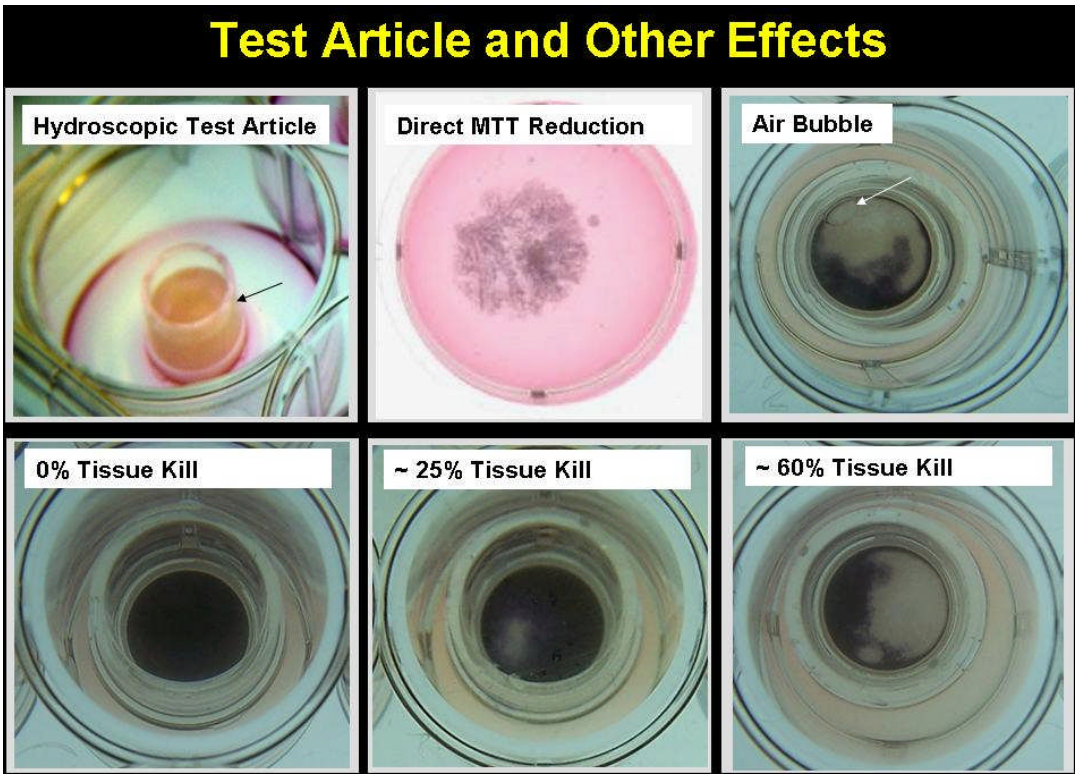


Figure 2-8 Photographs of various aspects of the EpiOcular assay.

A more complete description of the EpiOcular assay is given in a Background Review Document recently prepared for submission to ECVAM. Although the EpiOcular BRD focuses on a prediction model different from that proposed in this BRD, the treatment protocol is essentially identical, with the exception that the ECVAM BRD protocol uses a dilution of the test article before application. The ECVAM BRD will be referred to repeatedly in this submission where more detail is required.

2.2.3 Overview of how the BCOP test method is conducted

The overview of the BCOP test method procedures given below is taken directly from the NICEATM BRD “Current Status of *In vitro* Test Methods for

Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method."

"The basic procedures used to assess the effects of a test substance on an isolated bovine cornea were first reported by Gautheron *et al.* (1992). As described by Sina and Gautheron (1994, 1998), the BCOP assay uses isolated corneas from the eyes of freshly slaughtered cattle. Corneas free of defects are dissected with a 2 to 3 mm rim of sclera remaining to assist in subsequent handling, with care taken to avoid damage to the corneal epithelium and endothelium. Isolated corneas are mounted in specially designed corneal holders that consist of anterior and posterior compartments, which interface with the epithelial and endothelial sides of the cornea, respectively (Figure 2-9 – upper left). Both chambers are filled with medium and the device is then incubated at 32 ± 1°C for one hour to allow the corneas to equilibrate with the medium and to resume normal metabolic activity. Following the equilibration period, fresh medium is added to both chambers, and a baseline opacity measurement is performed. Corneal opacity is measured quantitatively as the amount of light transmission through the cornea (Figure 2-9 – upper right).

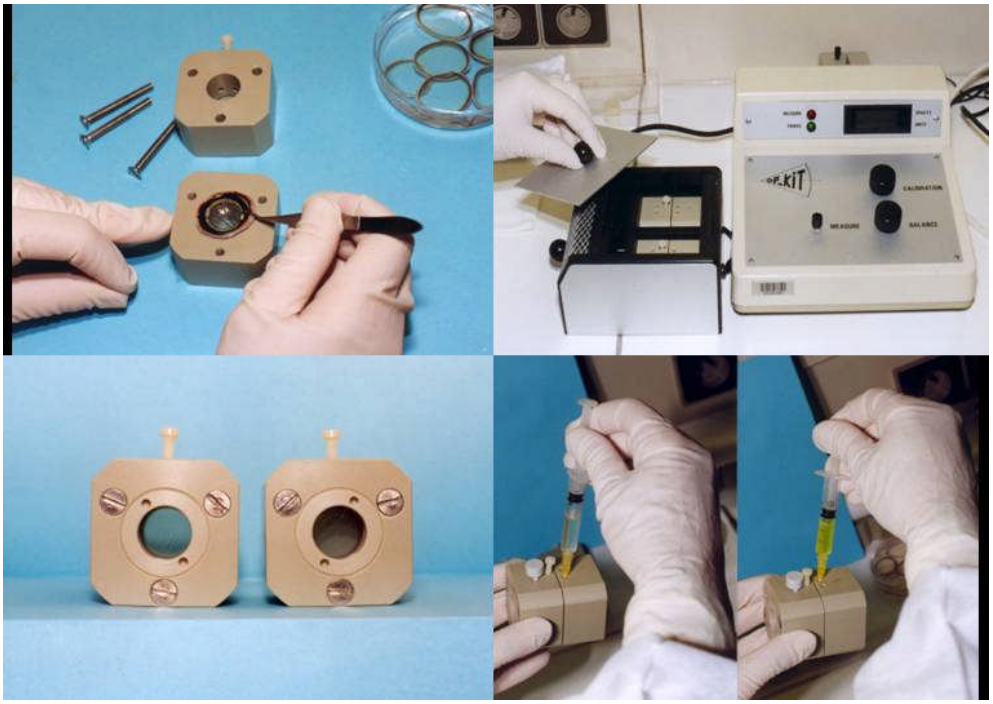


Figure 2-9 Photographs of various procedures occurring in the BCOP protocol. Upper left – Placing an excised cornea on the corneal holder. Upper right – Using the opacitometer to measure the opacity of a bovine cornea contained in a corneal holder. Bottom left – Visual comparison of the transparency of an untreated cornea on the left and a cornea treated with an irritating material on the right. Lower right – removing fluorescein solution from the posterior chamber prior to measuring its optical density in a spectrophotometer.

Two treatment protocols are used, one for liquids and surfactants, and one for solids. The protocol used by IIVS for the majority of the studies in this BRD is

given in Annex A4. Test substances are applied to the epithelial surface of the cornea by addition to the anterior chamber of the corneal holder.

The positive controls currently used for BCOP studies at IIVS are ethanol (neat) for the liquids protocol, and imidazole (200 mg/mL in complete MEM without phenol red) for the solids protocol. The current (as of 4/28/08) acceptable *In Vitro* Scores are 51.9 ± 6.2 for ethanol and 100.0 ± 15.9 for imidazole..

Liquids are tested undiluted; pure surfactants are generally tested at a concentration of 10% in saline or deionized water. Corneas are incubated horizontally for 10 ± 1 minutes at 32 ± 1 °C. The test substance is removed from the anterior compartment and the epithelial surface is washed at least three times. After refilling both chambers with fresh medium, a second opacity measurement is taken and the corneas are incubated again at 32 ± 1 °C for two hours prior to taking a final opacity measurement.

Solids are tested as solutions or suspensions at 20% concentration in saline or deionized water. Corneas are incubated horizontally for four hours at 32 ± 1°C. The test substance is removed from the compartment and the epithelial surface is washed at least three times with medium or until the corneal surface is free of visible particles. Fresh medium is added to both chambers and an opacity measurement is taken without further incubation.

Immediately after completing the final opacity measurements, corneal permeability is determined quantitatively by evaluating changes in the barrier properties of the epithelium to sodium fluorescein. To the anterior compartment of the corneal holder, 1 mL of sodium fluorescein (0.4% for liquids and surfactants, 0.5% for solids) is added. The corneas are incubated horizontally for 90 minutes at 32 ± 1°C. The amount of dye that penetrates the cornea is determined by measuring the OD of the medium in the posterior chamber (Figure 2-9 – lower right) with a microplate reader or UV/VIS spectrophotometer set at 490 nm.

A mean corrected opacity value (± standard deviation [SD]) and a mean corrected permeability value (OD units ± SD) are calculated for each treatment group. Most BCOP studies calculate an *In vitro* Score for irritancy that combines both values using the following empirically derived formula (Sina, Galer et al. 1995): *In vitro* Score = opacity value + 15 x OD_{490} value.

Generally, a substance producing an *In Vitro* Score from 0 to 25 is considered a mild irritant, from 25.1 to 75 (to 55 in early studies with pharmaceutical intermediates) a moderate irritant, and from 75.1 and above a severe irritant. A few laboratories do not calculate an *In Vitro* Score, but evaluate the opacity and permeability values independently. Also, some companies, such as S.C. Johnson & Son, Inc., do not use the classification system described above to assign an ocular irritancy classification, but instead compare BCOP data for newly tested substances to benchmark materials, relying on a system of comparative toxicity instead of cutoff

scores (Cuellar N and Swanson J, personal communication). In some cases, S.C. Johnson could also use a combination of classification scheme, control scores, histology, and knowledge about the chemistry of the formula to evaluate the test substance appropriately (Cuellar, N, personal communication).

These procedures were initially developed to assess the ocular irritation potential of pharmaceutical manufacturing intermediates and raw materials (Gautheron, Giroux et al. 1994; Sina 1994). However, as the BCOP test method gained more widespread use, the protocol has been modified by different investigators interested in using the assay to evaluate the ocular irritancy potential of other types of materials, including surfactant-based personal care cleaning formulations (Gettings, Lordo et al. 1996), home care products (Casterton, Potts et al. 1996), alkaline liquid laundry detergents (Cater, Nusair et al. 2002), oxidizing/reactive cleaning products (Swanson, White et al. 2003) and petrochemical products (Bailey, Freeman et al. 2004). As a result of the different testing needs of different investigators, additional endpoints have been used, such as assessment of corneal hydration (Ubels 1998; Cooper, Earl et al. 2001; Jones, Budynsky et al. 2001), and histological assessment of morphological alterations in the cornea (Curren, Evans et al. 2000; Swanson and Harbell 2000; Cater, Raabe et al. 2001; Cooper, Earl et al. 2001; Jones, Budynsky et al. 2001; Burdick, Merrill et al. 2002).

If a histological evaluation of the cornea is performed, the cornea is fixed in an appropriate fixative (*e.g.,* 10% neutral buffered formalin) after completing the corneal permeability steps of the assay. The cornea is fixed at room temperature for at least 24 hours before processing. After embedding the corneas, they are sectioned and stained with an appropriate stain such as hematoxylin and eosin. Corneal sections are examined for lesions in the epithelium, stroma, and endothelium. Sections from treated corneas are compared to those from concurrent negative and positive control corneas (Evans 1998; Curren, Evans et al. 2000)).

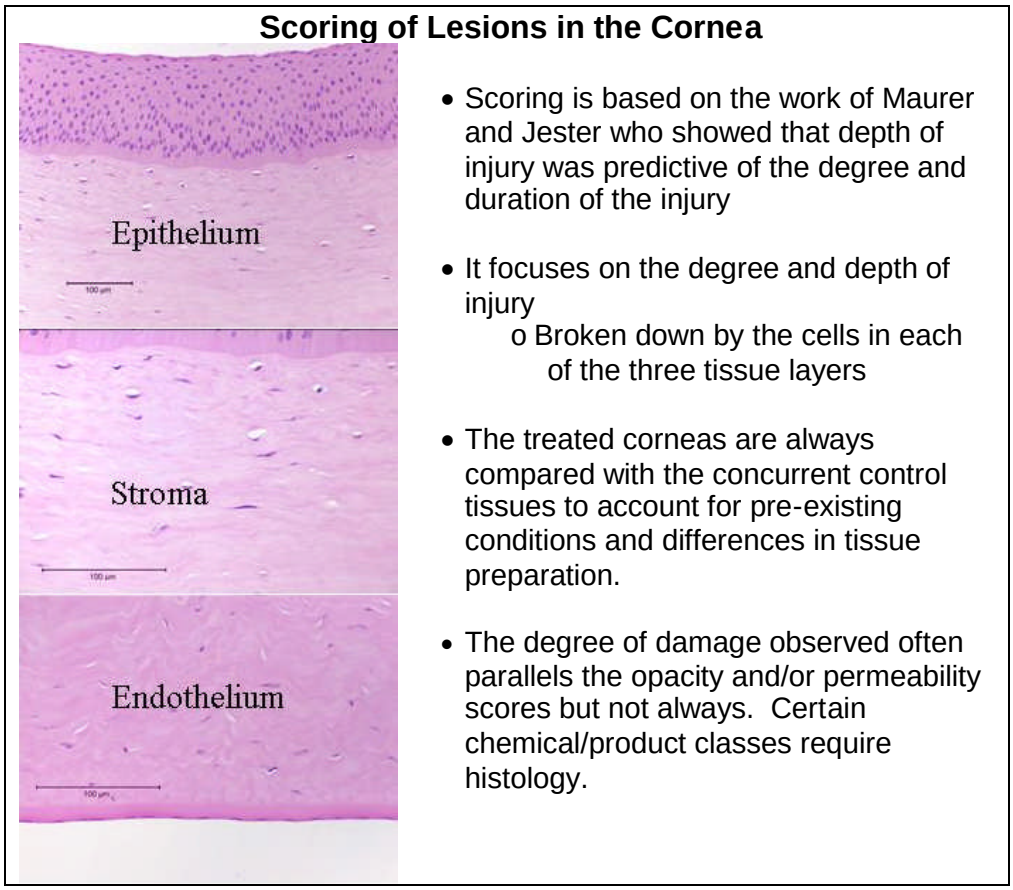


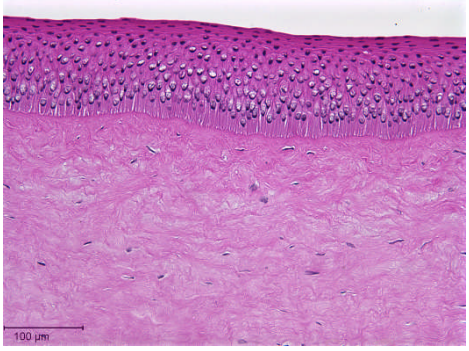
Figure 2-10 Histological evaluation of corneas

Other common modifications to the basic BCOP protocol include use of variable test substance exposure times and post-exposure periods that are specific to certain types of substances or products. For example, shorter exposure times are sometimes used for volatile organic solvents (Harbell J, personal communication; (Cuellar, Lloyd et al. 2003; Cuellar, Lloyd et al. 2004), longer exposure times are used for diluted materials or for increased sensitivity in the mild range of irritancy (Gettings, Lordo et al. 1996; Bruner, Carr et al. 1998; Cater, Nusair et al. 2002; Cater, Mun et al. 2003), and longer post-exposure expression periods are sometimes used to test substances with a potentially delayed onset of irritancy (Rees, Swanson et al. 2001; Cuellar, Lloyd et al. 2003; Gran, Swanson et al. 2003; Cuellar, Lloyd et al. 2004)."

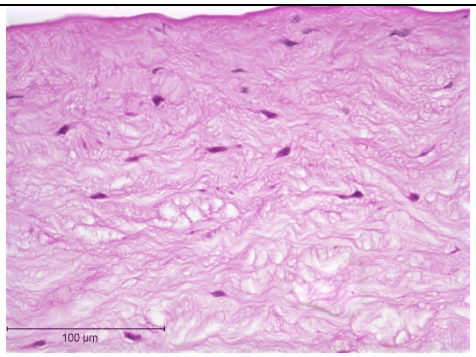
A more complete description of the BCOP assay is given in a Background Review Document prepared by NICEATM and amended by a Peer Review Panel. The BRD is attached as an annex to this submission and will be referred to repeatedly in this submission where more detail is required.

2.3 Use of histology in conjunction with the BCOP assay

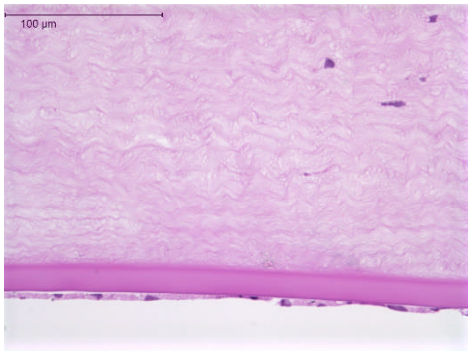
Histological evaluation of bovine corneas has been conducted at IIVS for approximately 8 years. During this time we have developed standard practice for the evaluations which have been consolidated into a guidebook (Annex G). The guidebook describes the process of evaluation and also contains a set of photomicrographs illustrating the various lesions that are found in treated corneas. This guidebook can be found in Annex G. A recent meeting (June 2008) of experts in ocular histopathology examined this document and will continue to work together to create a final consensus guidebook for the field. Figure 2-11 gives examples of epithelial damage, upper stroma damage, and lower stroma/endothelial damage.



(a) Epithelial damage with squamous layer coagulation and cytoplasmic and nuclear vacuolization in the wing and basal layers (20X).



(b) Severe collagen matrix vacuolization of the upper stroma. Note also the destruction of the upper keratocytes.



(c) Damage to the endothelial cell layer, cytoplasmic, and deep stromal collagen matrix vacuolization (severe).

Figure 2-11 Corneal damage after exposure to test article in the BCOP assay.

3 Substances Used For Validation of the Proposed Testing Approach

3.1 Rationale for the products selected, including rationale for solicitation of additional test materials to fill in gaps

The goal of this BRD is to present evidence that an *in vitro* testing strategy can provide for adequate protective labeling of a well-defined product category – anti-microbial cleaning products. Therefore, only this class of products (or products which have similar formulations) were used to determine the relationship between the results of the *in vitro* tests and the results from historical *in vivo* testing (Draize or LVET eye irritation test), *i.e.,* the relevance of the test. To do this, the manufacturers who participated in this program chose to submit data on products for which *in vitro* and *in vivo* data existed and in many cases also for products for which *in vitro* data only was available.

When considering the reproducibility of the assays; however, it seemed reasonable to utilize as much information as was available even though this information was derived from a wide range of products and ingredients. Thus, we incorporated reproducibility information for the three individual assays that was available in previously written BRD's even though some of these data were derived from products which did not fall into the anti-microbial cleaning product category.

3.2 Rationale for dividing substances into "buckets"

Anti-microbial cleaning products can be formulated with various types of chemistries. Some products – generally containing solvents or surfactants - clean by causing physical changes to the soil which allows the soil to be more easily removed from the surface. Other products clean by causing chemical changes to the soil. This can be accomplished by using strongly alkaline or acidic formulations, or by using extremely reactive formulations containing such ingredients as bleach, peroxides, or percarbonates.

Because there very likely could be different modes of action whereby these products could cause eye irritation, we thought it prudent at the beginning of the study to classify each anti-microbial cleaning product into one (or more) of five subcategories – solvents, oxidizers, acids, bases, or surfactants – depending on the specific formulation. In many cases a product might also be assigned to a second or third subcategory if more than one mode of action was suspected. We thought it possible that certain types of products might have to be handled differently as they progressed through an *in vitro* testing strategy.

The following chemical descriptors were used to characterize the different types of chemically-induced mechanisms associated with ocular irritation. These

were chosen based on existing information about the mechanisms of ocular irritation and the common types of formulation chemistries used in commercial and household cleaning products. The primary (and additional) categories were assigned by the company toxicologist(s) whose product was being evaluated in this program.

- Surfactants (SU) (*e.g.,* cationic, anionic, and nonionic with limited acid or alkaline activity)
- Acids (AC) (*e.g.,* with pH <4, especially where reserve acidity would contribute to the irritation potential)
- Alkaline (AL) products (bases) (*e.g.,* with pH >9, especially where reserve alkalinity would contribute to the irritation potential)
- Solvents (SO) (where organic solvents are expected to contribute to the irritancy potential (*e.g.,* alcohols, glycol ethers, etc.))
- Oxidizers (RC; Reactive chemistry) (formulations containing specific reactive chemicals, *e.g.,* hypochlorite, peroxide, percarbonate, oxygen bleaches, etc.)

As the results of our *in vitro/in vivo* comparisons became available we planned to look at each subcategory of cleaning products separately to see if they were possibly responsible for a greater number of overpredictions or underpredictions than the other subcategories. If not, then there would be no reason to treat individual subcategories in a special way, and all of anti-microbial cleaning products could progress through exactly the same *in vitro* testing scheme.

At the end of the study we concluded that only two types of chemical formulations should be assigned a special testing program. We recommend that Oxidizers, because they were often overpredicted by the CM and EO assays, should be tested only with the BCOP assay. We also recommend that formulations with “high solvent” concentrations (>5%) – if they are tested in the BCOP assay – should be tested with a three minute exposure time rather than the normal ten minute exposure time.

3.3 Rationale for number of substances included in the study

The number of substances included in this study was determined only by the number of formulations for which paired *in vivo* and *in vitro* data existed. After evaluating these data and constructing preliminary prediction models, we tested the prediction models by *in vitro* testing of either existing products or product reformulations which had previously been tested *in vivo* but not *in vitro*. There was no statistical basis for the number of substances; the number was only limited by availability of previously animal tested products which were relevant for this initiative. <u>No new animal testing was done for the purposes of this project</u>.

3.4 Chemicals or products evaluated

The anti-microbial cleaning products were broken down into six subcategories depending on the composition of their formulation: solvents, oxidizers, surfactants, acids, bases, or other. Table 3-1 gives the distribution of each subcategory of chemicals based on the *in vitro* assay system.

Table 3-1 Descriptive subcategory of products tested in the individual assays. Final graphs may contain fewer materials as final applicability domains were determined.

Paired *In vitro* & *In vivo* Data Sets				
Subcategory of cleaning products	**Number of substances tested per assay**			
	Cytosensor	**EpiOcular**	**BCOP**	**Total**
Solvents	18	10	12	39
Oxidizers	0	13	16	33
Surfactants	82	17	18	114
Acids	1	2	7	10
Bases	4	11	14	29
Other	-	2	1	3
Total	105	55	68	228

3.5 Coding procedures

The individual manufacturers who participated in this study stated that the *in vivo* testing was generally done by providing the testing laboratory a product coded by a system that they had developed in house. Often these products were accompanied by an MSDS that would have described in general terms their chemical composition.

The same type of coding was used for materials that had undergone *in vitro* testing before the start of this project. Products which underwent *in vitro* testing in the course of this project were coded by the manufacturer before shipping to IIVS. The materials were accompanied with MSDS's contained in sealed envelopes. In case of emergency the envelopes could be opened to obtain safety information. In all cases, the envelopes were not opened and the products decoded until after the *in vitro* testing. In the case of the BCOP assay, some products were decoded after the primary assay, but the identity of the materials was withheld from the individuals responsible for histopathological evaluation of the samples until after the evaluations were completed.

4 *In vivo* Reference data used for the assessment of accuracy

4.1 Protocols used to generate the *in vivo* data

4.1.1 Draize rabbit eye irritation protocol

The test method currently utilized for the majority of eye irritation tests conducted today, and also for the majority of *in vivo* eye irritation data presented in this BRD, is the Draize rabbit eye test. A good description of the Draize test is presented in the NICEATM BRD for the BCOP assay and is quoted directly below:

> "The methodology, originally described by Draize *et al.* (1944), involves instillation of 0.1 mL of the test substance (*e.g.,* liquids, solutions, and ointments) into the conjunctival sac of an albino rabbit eye. In this test method, one eye is treated while the other eye serves as the untreated control. The eye is examined at selected time intervals after exposure and any injuries to the cornea, conjunctiva, and the iris are scored. Scoring is subjective and based on a discrete, arbitrary scale (*reference omitted*) for grading the severity of ocular lesions. The scores for the observed ocular injuries range from 1 to 2 for iris effects, from 1 to 3 for conjunctival redness and discharge, and from 1 to 4 for corneal effects and conjunctival chemosis. A score of zero is assigned when the eye is normal and no adverse effects are observed. In the original protocol, the eyes were observed up to 4 days after application of the test substance. However, in current practice these time points vary according to the degree of irritation, the clearing time, and testing requirements imposed by the various regulatory agencies.
>
> The original Draize protocol describes a scoring system in which each ocular parameter is graded on a continuous numerical scale. The scores may be weighted (see Table 4-1); however, most classification systems today do not use a weighting factor. The weighting of the score by Draize *et al.* (1944) is biased more heavily for corneal injury, since injury to the cornea has the greatest probability of producing irreparable eye damage. To illustrate, each ocular parameter shown in (Table 4-1) is evaluated for each rabbit. The product of the opacity and area scores is obtained, then multiplied by a weighting factor of 5; the maximum corneal score is 80. The iris score is multiplied by a weighting factor of 5; the maximum score is 10. The scores for the three conjunctival parameters are added together and then the total is multiplied by a weighting factor of 2; the maximum score is 20. The overall score for each rabbit is calculated by adding the values for each parameter; the maximum total score is 110."

Table 4-1 Scale of weighted scores for grading the severity of ocular lesions (Draize, Woodard et al. 1944).

I. Cornea	
A. Opacity-Degree of density (area which is most dense is taken for reading)	
Scattered or diffuse area-details of iris clearly visible	1
Easily discernible translucent areas, details of iris slightly obscured	2
Opalescent areas, no details of iris visible, size of pupil barely discernible	3
Opaque, iris invisible	4
B. Area of cornea involved	
One quarter (or less), but not zero	1
Greater than one qu arter, but less than one -half	2
Greater than one-half, but less than three quarters	3
Greater than three quarters up to whole area	4
Score equals A x B x 5 Total maximum = 80	
II. Iris	
A. Values	
Folds above normal, congestion, swelling, circumcorneal injection (any one or all of these or combination of any thereof), iris still reacting to light (sluggish reaction is positive)	1
No reaction to light, hemorrhage; gross destruction (any one or all of these)	2
Score equals A x 5 Total possible maximum = 10	
III. Conjunctiva	
A. Redness (refers to palpebral conjunctiva only)	
Vessels definitely injected above normal	1
More diffuse, deeper crimson red, individual vessels not easily discernible	2
Diffuse beefy red	3
B. Chemosis	
Any swelling above normal (includes nictitating membrane)	1
Obvious swelling with partial eversion of the lids	2
Swelling with lids about half closed	3
Swelling with lids about half closed to completely closed	4
C. Discharge	
Any amount differ ent from normal (does not include small amount observed in inner canthus of normal animals)	1
Discharge with moistening of the lids and hairs just adjacent to the lids	2
Discharge with moistening of the lids and considerable area around the eye	3
Score equals (A + B + C) x 2 Total maximum = 20	

Although the above paragraph refers to the calculation of a numerical score to characterize eye irritation potential, the approach taken in this BRD is to translate individual tissue scores observed into toxicity categories, *e.g.,* the EPA toxicity categories or the GHS categories, which are described later (Sections 4.3 and 4.4).

A more detailed description of the Draize eye irritation method for observing and scoring tissue lesions, test guidelines for various international regulatory agencies, and other details of the test are given in the NICEATM BRD on the BCOP assay. In some cases a modified Draize procedure which utilized a 30 µl dose of

test material to the conjunctival sac was used (See section 4.2 Original reference data).

Despite the common use of the Draize eye irritation test it is not without its serious detractors (Daston and Freeberg 1991; Prinsen 2006).

4.1.2 LVET rabbit eye irritation protocol

The traditional Draize methodology described above has often been criticized for being very overpredictive of human response (Walker 1985). For example, 1) the amount of material (100 µL) dosed into the eye is more than the human eye, or even the rabbit eye can retain, 2) dosing in the conjunctival sac of the rabbit allows for much greater exposure to the test material than would the typical accidental exposure scenario to the human eye which would be a splash to the surface of the cornea, and 3) direct comparison of the human and rabbit ocular response to several types of cleaning products (Freeberg, Nixon et al. 1986; Roggeband, York et al. 2000) indicates that the rabbit response with the Draize protocol is much greater than that seen in the human.

In response to these concerns, a modification of the Draize eye irritation test – the Low Volume Eye Test (LVET) (Griffith, Nixon et al. 1980) – was developed and has been well characterized over a number of years. The essential difference is in dosing of the animals. In the LVET, a 10 µl dose is placed in the center of the cornea, in contrast to the traditional Draize methodology in which 100 µL is placed into the conjunctival sac. The LVET dosing regimen was to more closely model expected human exposure with a volume small enough that it could be retained in the eye. Scoring of the LVET is conducted identically to that of the Draize test according to the scale presented in Table 4.1.

The approach taken in this BRD is to translate the individual tissue scores observed into toxicity categories, *e.g.,* the EPA toxicity categories or the GHS categories, which are described later (Sections 4.3 and 4.4).

4.1.3 Comparison of Draize and LVET

It has been well reported that results obtained with the Draize eye irritation protocol (Draize, Woodard et al. 1944) do not reflect the eye irritation toxicity for humans. This was shown by the early work of Beckley (Beckley 1965; Beckley 1969). The rabbit Draize test grossly overpredicted the effects that you would see in the human eye (Lambert, Chambers et al. 1993).

The dose volume is one of the most influential factors that contribute to overprediction of the human response to detergent and cleaning products by the rabbit Draize test. The volume that is instilled into the lower conjunctival sac is 100µL, which exceeds the volume capacity of the rabbit eye lower conjunctival sac that can maximally hold ~80µL without blinking (Swanston 1985). The blink reflex is

also an important point. In the human, the spontaneous blink rate is about 12-20 per minute (Bell, Emslie-Smith et al. 1976; Karson, Berman et al. 1981) and serves to refresh the tear film at each blink. This is much more frequent than the spontaneous blink rate of about 3 blinks per hour in the rabbit (Mann and Pullinger 1942). Besides this spontaneous blinking, there is forced blinking in man in response to threat or injury. The blink reflex is a natural and involuntary response to a foreign material contacting the surface of the eye. Since the blink reflex is poorly developed in rabbits and highly developed in man, it is reasonable to take the blink reflex into account when considering the volume of a material that can contact the human eye. A volume of 100µL is approximately 10 times the normal volume of liquid (~10µL) residing in the human eye after blinking (Ehlers 1976; Swanston 1985). Equally important is that a volume of 100µL greatly exceeds (>10 times) the volume that directly covers the eye, *i.e.,* the tear volume of both the rabbit and the human eye (~ 7µL) (Mishima, Gasset et al. 1966; Chrai, Patton et al. 1973). Taking into account the anatomical facts, it is clear that the 10µL volume is more than the volume that can be in direct contact with either the rabbit or the human eye, *i.e.,* more than the tear volume.

The rabbit low volume eye test (LVET) addressed issues associated with the gross over-dosing and the animal welfare concerns of the Draize method (Griffith, Nixon et al. 1980). Correlation of recovery in the LVET with recovery in human accidents (Freeberg, Griffith et al. 1984; Freeberg, Hooker et al. 1986), and controlled comparative studies with 100µL and 10µL of detergent based products (Freeberg, Nixon et al. 1986), have shown that the LVET method is a better predictor than the Draize test, yet the LVET still overpredicts the human recovery time. Tables 4-2 and 4-3 summerize the results of the Freeberg *et al.* 1986 study where both rabbits and human volunteers (who were fully informed and participated in an Institutional Human Subjects Review Board-approved study) were exposed to identical concentrations of four representative household cleaning products. Table 4-2 shows that days-to-clear in the human were better predicted by the rabbit LVET assay than by the rabbit Draize assay, <u>although the rabbit LVET assay still overpredicted the effects of both the human 100 µL or 10µL exposure</u>. Table 4-3 extends this finding to the traditional Draize scoring scale. Again it can be seen that the rabbit LVET protocol predicts the human eye score better than the rabbit Draize protocol and <u>that the rabbit LVET protocol still overpredicts the effects of both the human 100 µL and 10 µL exposure</u>. Another example comes from Ghassemi *et al.* 1993 who compared the response of humans and rabbits to a liquid household cleaner (Table 4-4). By enumerating the number of eyes affected at the corneal, conjunctival or iridial level (or days-to-clear), it was again found that the rabbit LVET protocol overestimated the human reponse for all parameters with the exception of conjunctival involvement where it was equivalent.

Table 4-2 Mean time to clear after direct instillation of household cleaning products to both rabbits and humans. Compiled from Freeberg *et al.* 1986.

Product	Draize Protocol		LVET Protocol	
	Rabbit	Human	Rabbit	Human
Liquid fabric softener (100%)	3.5 days	12.5 hours	1.1 days	13.2 hours
Liquid shampoo (20%)	2.6 days	7.9 hours	1.4 days	7.5 hours
Liquid hand soap (10%)	2.7 days	9.1 hours	1.8 days	10.5 hours
Liquid laundry detergent (4%)	3.1 days	19.8 hours	1.7 days	4.8 hours

Table 4-3 Rabbit and human eye responses after exposure to either 100 µL (Draize protocol) or 10 µL (LVET protocol). All scoring done by the traditional Draize scoring scale. Compiled from Freeberg *et al.* (1986)

Reading time (hours)	Draize Protocol		LVET Protocol	
	Mean rabbit score	Mean human score	Mean rabbit score	Mean human score
Liquid Fabric softener (100%)				
1	4.3	0.8	4.8	1.8
24	6.5	-[a]	0.3	-
48	3.0	-	0.0	-
72	0.8	-	-	-
Liquid Shampoo (20%)				
1	11.1	4.0	6.0	2.0
24	7.0	-	0.8	-
48	4.3	-	0.0	-
72	0.9	-	-	-
Liquid hand soap (10%)				
1	8.0	3.0	4.0	2.5
24	13.9	-	1.8	-
48	4.3	-	0.3	-
72	0.3	-	0.3	-
Liquid laundry detergent (4%)				
1	8.3	4.0	4.5	2.3
24	13.3	-	1.8	0.0
48	9.0	-	0.5	-
72	1.4	-	0.0	-

[a]not scored

Table 4-4 Rabbit and human eye responses after exposure to either 100 µL (Draize protocol) or 10 µL (LVET protocol) for the liquid household cleaner. All scoring done by the tradititional Draize scoring scale. Compiled from Ghassemi *et al.* (1993)

Dosing Procedure	Number of Eyes Affected			Max. Time to Clear
	Cornea	Iris	Conjunctiva	
Rabbit LVET	3/3	2/3	3/3	7 days
Human LVET	0/10	0/10	10/10	2 days
Human Draize	0/10	0/10	10/10	< 3 days

In addition, comparisons can be made between predictions made by either the rabbit LVET or Draize test and human experience from accidental exposure (Freeburg et al 1986b). Table 4-5 shows that mean Time-to-Clear in days for these household cleaning products is always shorter in the human accidental exposure data than was predicted by either the Draize of LVET information. Additional

information exists on the overprediction of the LVET protocol (Bruner and Kohrman 1993; Cormier, Hunter et al. 1995), including an additional study directly comparing effects of low volumes of undiluted detergent and cleaning products in humans and rabbits (Ghassemi, Sauers et al. 1993; Roggeband, York et al. 2000).

Table 4-5 Average Time-to-Clear (days) for ocular effects following accidental exposure in humans and in rabbit eye irritation tests (LVET and Draize test) to household and cleaning products (Freeberg, Hooker et al. 1986).

Product[a]	Average Time-to-Clear (Days)		
	Human Data	LVET	Draize
Liquid Laundry Product #1	1.92	26.6	35
Liquid Dishwashing Product #1	0.77	8.2	25.7
Dry Dishwashing Product #1	0.59	4.6	18.3
Liquid Dishwashing Product #2	0.43	7.7	11.7
Liquid Household Cleaning Product #1	0.38	-	11.1
Liquid Dishwashing Product #3	0.3	3.9	22.2
Liquid Household Cleaning Product #2	0.23	4	15.2
Dry Household Cleaning Product #1	0.19	1.3	29.2
Dry Dishwashing Product #1	0.08	2.1	13.8
Dry Dishwashing Product #2	0.06	2.9	15.1

[a]Laundry Products: additives, main wash detergents, fabric softeners; Dishwashing products: automatic and hand detergents; Household Cleaning Products: hard surface cleaners, non-abrasive cleaners

4.2 Original reference data

Supporting animal data for the comparisons made in this BRD came from three basic methodologies: 1) the traditional Draize protocol utilizing 100 µL (or 100 mg) dose of test article into the conjunctival sac, 2) a modified Draize protocol which involved dosing with 30 µL (or 30 mg) of material into the conjunctival sac, and 3) the LVET which involves dosing with 10 µL directly onto the surface of the cornea.

In one case, animal data came from the EPA guideline for assessing aerosols. In this protocol, the animal eye was held open while a 1 second spray of the test article was directed onto the cornea. This one data point was then paired with data from a specially designed BCOP study in which the bovine cornea was exposed to a similar 1 second spray of the test material. Other aspects of the BCOP protocol remained the same.

Some of the animal data from the 30 µL Draize protocol could not be used for the comparisons found in this BRD. If the final classifications were less than an EPA Category I or less than a GHS Category 1, the data were not used since it could be assumed that a higher dose of test material might have resulted in a higher categorization. On the other hand, if the resulting score was an EPA Category I or a GHS Category 1, the data were used since the assumption was that a higher dose would not have resulted in a lower score. Seven materials are included in this BRD which had the 30 µL protocol and resulted in an EPA Category of 1, while only six

materials could be included in the GHS analysis since one of the seven materials had a GHS Category of 2A.

The actual animal data were supplied to IIVS in one of two ways; either as copies of the final reports from the organization that conducted the animal studies or as Excel© spreadsheets which contained the full tissue scores that had been entered by the staff of the submitter. The Excel© spreadsheets which were submitted are contained in Annex C. For reasons of confidentiality, copies of final reports that were submitted as the primary source for the animal scores are not included in this BRD; only spreadsheets containing the data transcribed by IIVS employees are appended. However, some of the final reports are available for inspection by NICEATM or EPA staff upon request.

4.3 Description of EPA toxicity categories

The EPA uses four toxicity categories which determine the labeling information for the product. Table 4-6 lists the four categories along with the ocular endpoints for determining the toxicity category.

Table 4-6 EPA Eye irritation toxicity categories (EPA 2003)

EPA Category	Draize Eye Test Scoring
Category I	- Corrosive, corneal involvement or irritation (iris or cornea score ≥ 1 or redness or chemosis ≥ 2) persisting more than 21 days or Corneal effects that are not expected to reverse by 21 days
Category II	- Corneal involvement or irritation clearing in 8-21 days
Category III	- Corneal involvement or irritation clearing in 7 days or less
Category IV	- Minimal or no effects clearing in less than 24 hours*

* Based on positive scores for conjunctival irritation ≥ 2

The eye irritation toxicity indicator is based on the outcome of the Draize eye test. In this BRD we have also classified the toxicity on the basis of the LVET. At least three animals are tested per chemical (a one-animal screen protocol is permitted to determine if the chemical is a severe irritant). The most severe response of the animals is used to calculate the EPA toxicity category. A single animal with a Category I response would lead to a Category I classification regardless of the outcome of the other animals. The criteria used to determine if a given animal result could be used for the analyses in this BRD are the same as were used by NICEATM in their BRD on the BCOP test and are quoted below:

- "At least three rabbits were tested in the study, unless a severe effect (*e.g.*, corrosion of the cornea) was noted in a single rabbit. In such cases, substance classification could proceed based on the effects observed in less than three rabbits.
- A volume of 0.1 mL or 0.1 g was tested in each rabbit. A study in which a lower quantity was applied to the eye was accepted for

substance classification, provided that a severe effect (*e.g.,* corrosion of the cornea, lesion persistence) was observed in a rabbit.

- Observations of the eye must have been made, at minimum, at 24-, 48-, and 72-hours following test substance application, if no severe effect was observed.
- Observations of the eye must have been made until reversibility was assessed, typically meaning that all endpoint scores were cleared. Results from a study terminated early were not used, unless the reason for the early termination was documented."

4.4 Description of GHS toxicity categories

The GHS (UN 2003) classification system for eye irritation is also utilized in this BRD because of the likelihood that EPA labeling decisions will eventually be made on the basis of this system. The classification system was applied to animal data in this BRD in an identical fashion to that used by NICEATM in their BRD on the BCOP assay. This methodology is described below in an extract from their BRD.

> "The classification of substances using the GHS classification system (UN 2003) was conducted sequentially. Initially, each rabbit tested was classified into one of four categories (Category 1, Category 2A, Category 2B, and nonirritant) based on the criteria outlined in Table 4-7. The criteria provided in this table are identical to those described in the GHS classification and labeling manual (UN 2003). Once all rabbits were categorized, the substance classification was determined based on the proportion of rabbits with a single irritancy category."

Table 4-7 Criteria for Classification of rabbits according to the GHS classification system

GHS Category	Rabbit Category Necessary for Classification
Category 1	Group A: • Effects in the cornea, iris, or conjunctiva that were not expected to reverse or did not fully reverse[1] within the observation period of 21 days, or • A corneal opacity score of 4 at any time during the test Group B: Rabbit with mean scores (averaging of the scores on day 1, 2, and 3) for opacity ≥3 and/or iritis ≥1.5
Category 2A	Rabbit with mean scores (rabbit values are averaged across observation days 1, 2, and 3) for one or more of the following: • 1 ≤ Iritis < 1.5 • 1 ≤ Corneal opacity < 3 • Redness ≥ 2 • Chemosis ≥ 2

	and the effects fully reverse within 21 days
Category 2B	Rabbit with mean scores (rabbit values are averaged across observation days 1, 2, and 3) for one or more of the following: • 1 ≤ Iritis < 1.5 • 1 ≤ Corneal opacity < 3 • Redness ≥ 2 • Chemosis ≥ 2 and the effects fully reverse within 7 days
Nonirritant	Rabbit mean scores fall below threshold values for Category 1, 2A, and 2B

[1]Full reversal of the effects was defined as corneal, iritis, redness, and chemosis = 0.

After each rabbit was categorized, the ocular irritancy potential of the substance was determined. As shown in Table 4-8, substance classification depended on the proportion of rabbits that produced the same response. As noted above, if a substance was tested in more than three rabbits, decision criteria were expanded. Generally, the proportionality needed for classification was maintained (*e.g.,* 1 out of 3 or 2 out 6 rabbits were required for classification for most categories). However, in some cases, additional classification rules were necessary to include the available data. These additional rules are distinguished by italicized text in Table 4-8.

If an unequivocal substance classification could not be made due to the response pattern of the tested rabbits for a substance (*e.g.,* one rabbit classified as Category 1, Group B; two rabbits classified as Category 2B; three rabbits classified as nonirritant), the data were not used in the analysis.

Table 4-8 Criteria for Classification of Substance According to the GHS Classification System (Modified from UN 2003)

GHS Category	Criteria Necessary for Substance Classification
Category 1	1. At least 1 of 3 rabbits or 2 of 6 rabbits classified as Category 1, Group A *2. One of six rabbits classified as Category 1, Group A and at least 1 of 6 rabbits classified as Category 1, Group B* 3. At least 2 of 3 rabbits or 4 of 6 rabbits classified as Category 1, Group B
Category 2A	1. At least 2 of 3 rabbits or 4 of 6 rabbits classified as Category 2A *2. One of 3 (2 of 6) rabbits classified as Category 2A and 1 of 3 (2 of 6) rabbits classified as Category 2B*
Category 2B	1. At least 2 of 3 rabbits or 4 of 6 rabbits classified as Category 2B
Nonirritant	1. At least 2 of 3 rabbits or 4 of 6 rabbits classified as nonirritant

Italicized text indicates rules that were developed to include additional data.

4.5 Transformation of original data to toxicity categories

To transform the original data – existing either as a submitted spreadsheet or as an original report from the laboratory conducting the Draize eye irritation test – individual eye scores were entered into Excel® spreadsheets designed to categorize the scores according to the above listed criteria. Example spreadsheets can be found in Annex C. The spreadsheet used to determine EPA toxicity categories was designed at IIVS, and the spreadsheet used to determine GHS toxicity classifications was designed at ECVAM and supplied to IIVS.

As part of our additional analysis of the EPA and GHS classifications, information from six rabbit Draize tests was entered into a secondary spreadsheet which calculated GHS and EPA categories for each of the 20 distinct sets of 3 rabbit combinations as described in Section 4.8.

4.6 Quality of *in vivo* data

It is the generally stated goal of most validation authorities that data submitted in support of a validation effort should be conducted to comply with GLP guidelines (ICCVAM 1997; Hartung, Bremer et al. 2004). The GLP-compliance status of the majority of the animal studies in this BRD is not known since that information was not supplied by the sponsors. However, for a minority of the data the actual study reports were available, and it could be determined from these reports whether or not the studies were GLP-compliant. In cases where the studies were determined to be GLP-compliant this fact was noted in the spreadsheets.

4.7 Human toxicity information on cleaning products

We have no human toxicity information for any of the specific materials that are used as references for the *in vitro* results in this BRD. However, data do exist in the literature for certain types of cleaning products. Although it is not routine, ethically designed human studies have been conducted on such products. In addition, human accidental exposure data have been collected for some household cleaning products and this information compared with data from the Draize eye irritation test, the LVET, and human clinical studies. Several of these studies have already been discussed in detail in Section 4.1.2 of this BRD.

Essentially, the data indicate that the results of both the Draize test and the LVET overpredict the amount of damage that would occur in the human eye; however, the Draize test overpredicts by a greater amount.

4.7.1 Clinical Studies by Beckley *et al.* (1965) on a light duty liquid detergent

Beckley *et al.* (1965) compared a light duty liquid detergent (Table 4-9) on the eyes of rabbits, dogs, monkeys and humans (Beckley 1965).

Table 4-9 Composition of the light duty liquid detergent from the Beckley 1965 study (Beckley 1965)

Test Product	Ingredients	Level in Product (%)	Concentration Tested
Light Duty Liquid Detergent (pH 6.3)	Alkylbenzene sulphonate Conventional organic foam builder and solubilizer	38%	Various amounts, up to and including undiluted material
	Ethyl alcohol	12%	
	Water	50%	

The laboratory animal studies showed clear differences between species with the most sensitive being the rabbit, followed by the dog, and finally by the monkey. All of the animals whose eyes were not flushed showed some corneal involvement. Extracted results from the manuscript are shown in Table 4-10.

Table 4-10 Mean Draize scores for individual ocular tissues of six rabbits, six dogs and four monkeys (unflushed) or three animals each (flushed) after instillation of 100 µL of a Light Duty Liquid Detergent (Beckley 1965)

Evaluation Time	Ocular Tissue	Eyes Unflushed			Eyes Flushed		
		Rabbit	Dog	Monkey	Rabbit	Dog	Monkey
1h	Cornea	33.3	40.0	20.0	15.0	40.0	0
	Iris	10.0	5.0	2.5	10.0	0	0
	Conjunctiva	12.0	4.0	1.0	10.0	0	0
1 day	Cornea	33.3	45.0	10.0	13.3	20.0	0
	Iris	10.0	5.0	0	10.0	0	0
	Conjunctiva	12.0	4.0	0	10.0	0	0
3 days	Cornea	21.7	30.0	0	5.0	20.0	0
	Iris	10.0	5.0	0	8.3	0	0
	Conjunctiva	9.3	0	0	7.3	0	0
7 days	Cornea	6.7	0	0	1.7	0	0
	Iris	8.3	0	0	3.3	0	0
	Conjunctiva	6.7	0	0	2.7	0	0

In contrast to the animal results, three different studies using <u>human volunteers showed much milder reactions and no corneal involvement.</u>

- Study A: This study began with the instillation (100 μL) of increasing concentrations of the Light Duty Liquid Detergent into the lower conjunctival sac without rinsing. After it was determined that all of the diluted solutions were tolerated, undiluted solution was then instilled (100 μL) into the eyes of 15 volunteers for seven consecutive days. Ten of the subjects had no eye damage; five had began to develop conjuctivis which disappeared when dosing was stopped. There were no instances of corneal or iridial involvement.

- Study B: This study was an extended dosing study, again using fifteen subjects. It began with 100 μL instillations of increasing concentrations alternating daily between the left and right eye until the undiluted solution was used. Since 100 μL flooded the eye, the dosage was held in place for two minutes with a gauze pad. After removing the pad the eyes were rinsed. After the 20th day 100 μL of the undiluted solution was instilled into the same conjunctival sac for 8 days. A few subjects developed a conjunctival erythema. There were no instances of corneal or iridial involvement throughout the entire study.

- Study C: This study involved instilling three drops of undiluted solution into each eye for three days. "None of the subjects developed a chronic conjunctivitis, and in no case was there involvement of the iris or cornea."

The conclusion from this study is that humans are not only much less sensitive to this type of cleaning product than the rabbit, but also less sensitive than the dog and monkey.

4.7.2 Clinical Studies by Beckley *et al.* (1969) on a soap suspension and a liquid household cleaner

Beckley *et al.* (1969) also compared the effects of a 5% soap solution and an undiluted all-purpose liquid household cleaner on the eyes of rabbits, monkeys, and man. The composition of the all-purpose household cleaner is provided in Table 4-11.

Table 4-11 Composition of the test materials from the Beckley 1969 study (Beckley 1969)

Test Product	Ingredients	Level in Product (%)	Concentration Tested
Soap suspension	Soap	N/A	5%
Liquid Household Cleaner (pH 10.4)	Alkylbenzene sulphonate	5	Undiluted
	Ammonium cumene sulphonate	4	
	Builder containing 3% sodium carbonate and 1% tetrapotassium pyrophosphate	13	
	Miscellaneous ingredients including 0.7% ammonia, 0.6% soap and 0.4% perfume	2	
	Water	Up to 100	

- 5% Soap solution – Rabbits and monkeys had some corneal involvement lasting up to 72 hours for some of the animals. Both species had conjunctivitis up to 48 hours. The humans had some initial epithelial loss that was not observable at 6 hours. Conjunctivitis was seen in the humans at six hours but further measurements were not made.

- Liquid Household Cleaner – Rabbits corneal stippling up through 7 days and conjunctivitis through 3 days. Monkeys had corneal stippling through seven days and conjunctivitis up to 24 hours. Humans had corneal stippling only through 6 hours and conjunctivitis through 3 days.

The conclusion from this study was that humans had a slight corneal response to both 5% soap and the Liquid Household Cleaner, but it cleared by six hours. The laboratory animals, in contrast, had more severe responses.

4.7.3 Clinical Studies by Ghassemi, *et al.* (1997) on a liquid household cleaner

Ghassemi, *et al.* carried out direct installation studies in human volunteers with a liquid household cleaner of low pH. Table of 4-12 gives the composition of the cleaner.

Table 4-12 Liquid Household Cleaner composition used in the Ghassemi *et al.* (1997) study

Test Material	Ingredients	Ingredient Concentration	Concentration Tested
Liquid Household Cleaner (pH 3)	Nonionic surfactant: • alcohol ethoxylate	2%	Undiluted
	Amphoteric surfactant: • betaine	2%	
	• Na H_2 citrate	3%	
	• Cumene sulphonate	3%	
	Solvent: • butoxypropoxypropanol/ dipropylene glycol • monobutyl ether • Water	8% to 100%	

Undiluted Liquid Household Cleaner was instilled into one eye of ten human volunteers using either the Draize methodology (100 μL instillations) or the LVET methodlogy (10 μL onto the cornea). Rabbits were also dosed with the cleaner using the LVET method. There was no corneal or iridial involvement in the humans with either dosing procedure, but there was initial conjunctivitis which cleared by 48 hours after the 10 μL exposure and 70 hours after the 100 μL exposure. In contrast, the three rabbits had both corneal (3/3) and iridial (2/3) involvement, along with conjunctivitis which did not resolve until seven days.

The conclusion from this study is that human eyes are not significantly affected by this Liquid Household Cleaner (even with the 100 μL dosing volume), but rabbits have significant ocular responses to even the LVET procedure.

4.7.4 Clinical studies of liquid detergent products by Roggeband, *et al.* (2000)

Roggeband *et al.* conducted human clinical studies on two representative, surfactant-based cleaning products which are described in Table 4-13.

Table 4-13 Composition of the test materials from the Roggeband, *et al.* (2000) study

Test Product	Ingredients	Level in Product (%)	Concentration Tested
Concentrated Laundry Liquid	Soap	15	Undiluted
	Nonionic surfactant	27	
	Anionic surfactant	12	
	Water	Up to 100	
Concentrated Dishwasher Liquid	Non-ionic surfactant	4	Undiluted
	Anionic surfactant	38	
	Water	Up to 100	

Initial studies with the two test materials focused on finding dosing volumes of the two concentrated products that were just below the doses causing some corneal erosion. These doses were 3 μL for the Concentrated Laundry Liquid and 1 μL for the Concentrated Dishwashing Liquid.

Subsequently 10 human volunteers and six rabbits were exposed to identical doses of the Concentrated Laundry Liquid (3 μL) and the Concentrated Dishwashing Liquid (1 μL). Table 4-14 shows the results with the laundry liquid. At 1 hour in the human there were corneal effects in two volunteers, but there were no corneal lesions at 24 hours. There were also conjunctival effects at 1 hour, but these resolved in all but two volunteers at 24 hours. In the rabbit; however, there were corneal effects in 5 of the 6 rabbits at 24 hours, and rather strong conjunctival effects in all rabbits at 24 hours.

Table 4-14 Ocular responses of humans and rabbits to identical volumes (3 μL) of Concentrated Laundry Liquid. Modified from Roggeband, *et al* (2000).

Human					Rabbit				
	1 Hr		24 Hr			1 Hr		24 Hr	
Volunteer	Cornea[a]	Conjunctiva[b]	Cornea	Conjunctiva	Animal	Cornea	Conjunctiva	Cornea	Conjunctiva
A	0	1/1	0	0/0	A	0/0	1/1/0	1/2	2/1/1
B	0	1/0	0	0/0	B	0/0	1/1/0	1/2	2/1/1
C	0	1/0	0	0/0	C	0/0	1/1/0	0/0	2/1/1
D	1/2	1/0	0	1/0	D	0/0	1/1/0	1/4	2/1/0
E	1/1	1/0	0	0/0	E	0/0	1/1/0	1/3	2/1/1
F	0	1/0	0	1/0	F	0/0	1/1/0	1/4	2/1/1
G	0	1/0	0	0/0					
H	0	0/0	0	0/0					
I	0	1/0	0	0/0					
J	0	1/0	0	0/0					

[a]Corneal score expressed as opacity score/area
[b]Conjunctival score expressed as erythema score/edema score in humans and erythema/edema/discharge in rabbits.

The results with the dishwashing liquid are shown in Table 4-15. One of the human subjects had corneal involvement at 1 hour but this resolved by 24 hours. Three of the ten volunteers had a slight conjunctivial response at 1 hour, but all had resolved at 24 hours. In contrast 5 of the six rabbits had corneal opacities at 24 hours and all of the rabbits had conjunctival involvement at both 1 hour and 24 hours.

Table 4-15 Ocular responses of humans and rabbits to identical volumes (1 μL) of Concentrated Dishwshing Liquid. Modified from Roggeband, *et a* (2000).

Human					Rabbit				
	1 Hr		24 Hr			1 Hr		24 Hr	
Volunteer	Cornea[a]	Conjunctiva[b]	Cornea	Conjunctiva	Animal	Cornea	Conjunctiva	Cornea	Conjunctiva
A	0	0/0	0	0/0	A	0/0	1/1/0	0/0	1/1/1
B	0	0/0	0	0/0	B	0/0	2/1/0	1/2	2/1/0
C	0	0/0	0	0/0	C	0/0	1/1/0	1/1	2/1/0
D	1/1	1/0	0	0/0	D	0/0	1/1/0	1/1	2/1/0
E	0	0/0	0	0/0	E	0/0	1/1/0	1/2	2/1/0
F	0	0/0	0	0/0	F	0/0	1/1/0	1/2	2/1/0
G	0	1/0	0	0/0					
H	0	0/0	0	0/0					
I	0	1/0	0	0/0					
J	0	0/0	0	0/0					

[a]Corneal score expressed as opacity score/area
[b]Conjunctival score expressed as erythema score/edema score in humans and erythema/edema/discharge in rabbits.

The conclusions of this study were that concentrated surfactant cleaning products are capable of causing ocular effects in both the human and the rabbit. However, the effects in the rabbit after an identical dose to that applied to the human volunteers were more severe and resolved much later (some between 72 hr and seven days).

4.8 Accuracy and reliability of the LVET and Draize tests

A significant problem in analyzing how well any *in vitro* test predicts the outcome of an *in vivo* test is that a single value (without any estimate of error) is generally associated with the animal score for a test material, and this single value is treated as a "gold standard". In reality, there is no single eye irritation value that characterizes a test material; the value that is obtained will generally vary each time the material is tested. Thus, it is extremely unlikely that an *in vitro* score and an *in vivo* score will match exactly, no matter how perfectly the *in vitro* test is performed. This fact is often overlooked in most validation studies. Generally the animal score is treated as a single fixed value (since the animal test is generally conducted only once), and the *in vitro* test is then assessed for its "accuracy" based on how well its data match that of the animal test. Only a few studies, *e.g.,* the CTFA Phase III eye irritation evaluation study (Gettings, Lordo et al. 1996), have taken the animal test variability into account. The CTFA study used bootstrap resampling to estimate within group variability for each test material so that Draize scores could be represented more realistically with their variability (see, for example, Figure 4-1).

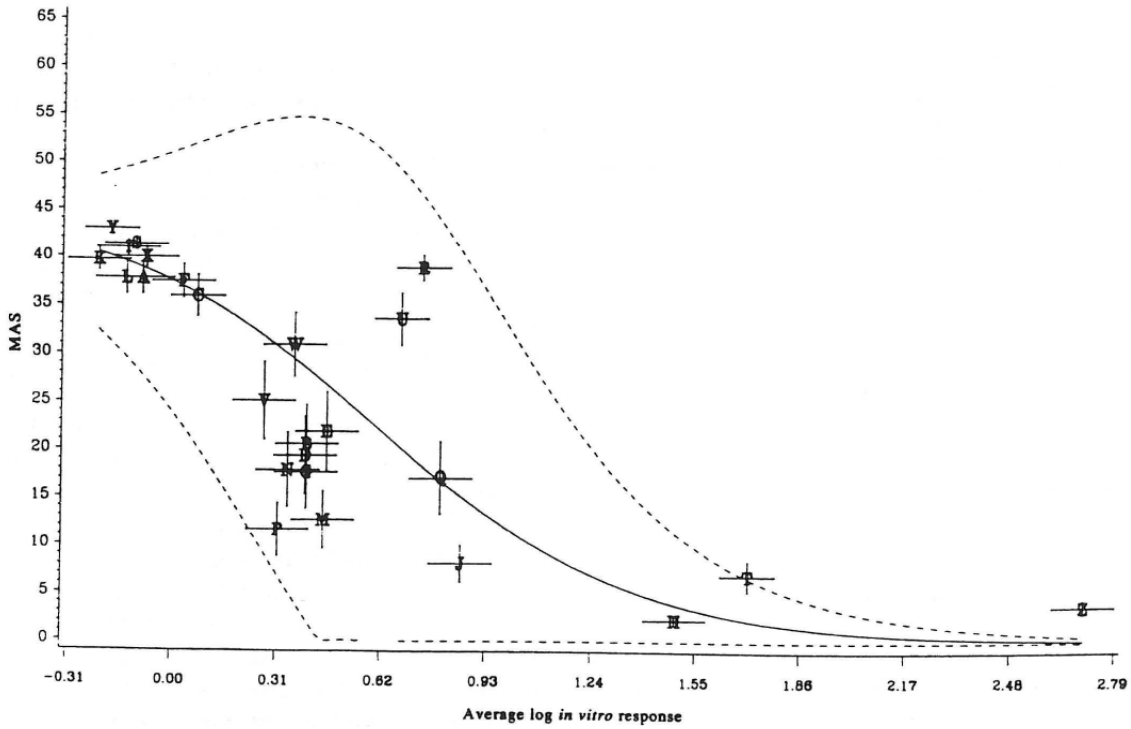


Figure 4-1. Performance of the Silicon Microphysiometer in predicting the Draize MAS score for test materials from the CTFA Phase III study of surfactant-based formulations (Gettings, Lordo et al. 1996). The variability associated with both the animal test and the *in vitro* test is shown on the graph.

4.8.1 Analysis of six rabbit tests in combinations of three

As mentioned above, one reason that Draize MAS scores are usually treated as unvarying values is that both ethical and financial considerations generally demand that a rabbit eye test only be conducted a single time. Thus for many materials there is no information about what score might occur in a repeat test, and without the results of multiple tests it is difficult to address variability.

However, there is one approach which can supply some quantitative insight into this problem. Because over the years the Draize test protocol has evolved from a six rabbit test to a three rabbit test, there is one way of estimating variability for materials which were tested with the six rabbit protocol. It is possible to analyze the ocular response of the six rabbits by placing them into smaller groups. For example, the results for each of the six individual rabbits can be recombined into multiple unique groups of three rabbits (matching the number of rabbits used in today's standard protocol). In fact, all rabbits (designated A – F in the following example) in a six rabbit test can be recombined into 20 unique three rabbit groups, e.g. ABC, ABD, ABE, ABF, etc. This is an approach already used by others in studies to determine the necessary sample size for a rabbit ocular irritation test (DeSousa, Rouse et al. 1984). Each three rabbit group can then be given a hazard classification according to the published guidelines from specific regulatory bodies. The number of subgroups in each hazard classification can then be viewed as a measure of the variability of the test. If all 20 subgroups are classified as R36, for example, then the R36 classification for that material can be considered not very variable. However, if 10 subgroups are rated as No Label and the other 10 are rated as R41, then the results for that material would be considered quite variable. In essence the above results mean that if the material were tested in multiple three rabbit tests, half of the tests would rate it as a very severe R41 material, and the other half of the tests would rate it as a mild No Label material. Therefore, an *in vitro* test of the same material should not necessarily be expected to always make a prediction of R41, which would be the overall prediction of the six rabbit test.

To demonstrate the level of Draize test variability which occurs in the real world, we have examined the animal data from the CTFA Phase III study. This study had arguably one of the best controlled animal studies because it was conducted under GLP's and utilized a randomized block design (3 males and 3 females) with each animal's dosing initiated on a separate day.

Table 4.16 shows for the CTFA Phase III study the number of three rabbit subgroups which fall into each of the hazard categories for the three regulatory classification schemes (GHS, EU, and EPA). Data which support these classifications can be found in spreadsheets contained in Annex C; CTFA Animal Data) It can be seen that in some cases all of the three rabbit subgroups give the same hazard classification as the six rabbit study, e.g. the EU classification for HZB, HZC and HZD is No Label, and each of the 20 three rabbit subgroups for each test material is also No Label. However, for those same three test materials classified by

GHS criteria there is considerable difference between the subgroups and the original six rabbit study. For example, HZC is No Label by the six rabbit test, but only half (10) of the three rabbit groups are No Label; seven are 2B and 3 are category 1. This means if the test were repeated 20 times using the current three rabbit protocol there would be an equal chance of having a higher than No Label score (10 out of 20 times) as there would be of having the No Label score (10 out of 20 times). Similar results can be seen for many of the materials in this study.

Even more dramatic examples can be found in the CTFA Phase III study. HZE, for example, is classified R41 by the six rabbit test, but only 10 of the subgroups have R41 classifications, the other 10 are No Label! Thus if the three rabbit test were run only once, there would be a 50% chance of having the lowest classification (No Label) and an equal chance of having the highest label (R41). HZP is another interesting example. Although it has a 6-rabbit GHS classification of No Label, 6 out of 20 tests (30% of the time) give a Category 1 result – three categories higher than that determined by the 6 rabbit test! Other interesting examples are highlighted in bold in the table.

Table 4-16 Recombination of each 6 rabbit test result into 20 three rabbit test subgroups. Each subgroup was classified separately according to the rules for each of the three classification systems, and the number of subgroups falling into each hazard category is indicated. Numbers in bold, shaded areas represent results from test materials where the subgroups differed in their hazard classification from the overall six rabbit classification. Data from the CTFA Phase III study. N = 25 materials.

		6 animal study score			GHS Counts				EU Counts			EPA Counts			
		GHS	EU	EPA	1	2A	2B	NL	R41	R36	NL	I	II	III	IV
Shampoo 7	HZA	1	R41	1	**16**	**4**	0	0	**16**	**3**	**1**	**16**	**4**	0	0
Liquid Soap 1	HZB*	NL	NL	3	0	0	**4**	**16**	0	0	20	0	0	20	0
Shampoo 1	HZC*	NL	NL	3	**0**	**0**	**10**	**10**	0	0	20	0	0	20	0
Shampoo 5	HZD*	NL	NL	3	0	0	0	20	0	0	20	0	0	20	0
Gel Cleaner	HZE	NL	R41	1	**10**	**0**	**0**	**10**	**10**	**0**	**10**	**10**	**0**	**10**	0
Baby Shampoo 2	HZF	1	R41	1	**16**	**4**	0	0	**16**	**3**	**1**	**16**	**4**	0	0
Shampoo 8	HZG*	NL	NL	3	**0**	**0**	**0**	**20**	0	0	20	0	0	20	0
Eye Makeup re.	HZH	NL	NL	4	0	0	0	20	0	0	20	0	0	0	20
Skin Cleaner	HZI	1	R41	1	**19**	**1**	0	0	**19**	**1**	0	**19**	**1**	0	0
Mild Shampoo	HZJ	NL	NL	4	0	0	0	20	0	0	20	0	0	0	20
Bubble bath	HZK	1	R41	1	20	0	0	0	20	0	0	20	0	0	0
Foam Bath	HZL	1	R41	1	**19**	**0**	**1**	0	**19**	**0**	**1**	**19**	**0**	**1**	0
Shampoo 3	HZM*	NL	NL	3	0	0	0	20	0	0	20	**0**	**0**	**10**	**10**
Shampoo 6	HZN*	NL	NL	3	0	0	0	20	0	0	20	0	0	20	0
Baby Shampoo 1	HZP	NL	NL	3	**0**	**0**	**0**	**20**	0	0	20	0	0	**19**	**1**
Cleaning Gel	HZQ	NL	NL	3	**0**	**0**	**0**	**20**	0	0	20	0	0	20	0
Facial Cleaning Foan	HZR*	NL	R41	1	**10**	**0**	**3**	**7**	**10**	**0**	**10**	**10**	**0**	**10**	0
Shower Gel	HZS	1	R41	1	**19**	**1**	**0**	0	**19**	**1**	0	**19**	**1**	0	0
Polishing Scrub	HZT	NL	NL	4	0	0	0	20	0	0	20	0	0	0	20
Hand Soap	HZU*	NL	NL	3	**0**	**0**	**4**	**16**	0	0	20	0	0	20	0
Shampoo 4	HZV*	NL	NL	3	0	0	0	20	0	0	20	0	0	20	0
Liquid Soap 2	HZW*	2B	NL	3	**0**	**0**	**16**	**4**	0	0	20	0	0	20	0
Shampoo 2	HZX	1	R41	1	**19**	**1**	0	0	**19**	**0**	**1**	**16**	**4**	0	0
Shampoo AntiD	HZY	1	R41	1	**16**	**4**	0	0	**16**	**4**	0	**16**	**4**	0	0
Facial Cleaner	HZZ	NL	NL	4	0	0	0	20	0	0	20	0	0	0	20

* tested at 25% (w/v) in vivo and in vitro (starting material)

The main conclusion from studying this example is that neither a Draize MAS score nor a Draize-defined EPA toxicity classification is an unvarying physical constant for the test material. Therefore, an *in vitro* test should not be expected to exactly match a toxicity category determined *in vivo* because the next time the animal test is run it might also fail to match the toxicity classification of the first animal test.

One other interesting piece of information can be found in the results in Table 4-16, and that is the EPA toxicity categories which would be assigned to this list of personal care and cosmetics products. The usual assumption is that EPA Category I materials are extremely toxic, such as undiluted commercial pesticides, or strong bleaches or acids. However, here we see that common products that are used routinely around the head and face are able to elicit Category I classifications. Even a labeled baby shampoo is categorized as an EPA Category 1! It is possible that many of these personal care products are actually potential severe eye irritants for

humans. However, this does not seem likely, or we would have seen a tremendous number of severe eye injuries from misuse (or even correct use) of the products. A more likely possibility is that the EPA scoring scale is quite overprotective of the human response. This is an important concept to keep in mind when assessing the predictive capacity of the *in vitro* tests described in this BRD. When assessing the validity of a new method it is always necessary to make some judgment concerning just how many underpredictions of the Draize-defined toxicity classifications can be accepted. Knowing how this set of personal care products scored in the Draize eye irritation test may assist in making realistic assessments.

4.8.2 Historic references on reliability of the Draize test

Additional information addressing the variability inherent in the Draize test can be found in (Weil and Scala 1971; Marzulli and Ruggles 1973; Choksi, Haseman et al. 2005; Prinsen 2006).

5 Test method data and results

Since the testing strategy described in this BRD consists of three separate test methods, the Cytosensor method, the EpiOcular method and the BCOP test method, information concerning the data and the protocols used to generate the data will be described under the appropriate headings for each test method in turn.

5.1 Description of the test method protocols used to generate data

The number of unique materials with *in vivo* and *in vitro* paired data is described by Table 5-1 for each assay system. The materials tested in the Cytosensor assay were not tested in any other *in vitro* assay system. The CTFA cytosensor study used the same 25 unique materials in both the Draize and LVET *in vivo* systems. The Colipa study and the CTFA study had some overlap of materials, but the materials were either reformulated or separately sourced with several year's time between the studies – thus it would be questionable to consider them "identical" materials. Thirty unique materials were tested in both the EpiOcular and BCOP assay systems. These materials are listed under EpiOcular, BCOP, and the EpiOcular & BCOP assays below.

Table 5-1 Description of number of unique materials tested in each assay system with corresponding *in vivo* data.

Assay	Study	*In Vivo* Data	Materials	Comments
Cytosensor	LVET Only Section 6.1.1	LVET	105 unique	Not tested in any other *in vitro* assay.
	CTFA Phase III Section 6.1.2.1	Draize LVET	25 unique 25 unique	Same 25 materials were tested in the Draize and LVET. Not tested in any other *in vitro* assay.
	COLIPA Section 6.1.2.2	Draize	20 unique	Not tested in any other *in vitro* assay.
EpiOcular	Different Companies Section 6.2.1	Draize LVET	30 unique 25 unique	Different materials tested in the Draize and LVET. 30 materials (all from the Draize study) were also tested in the BCOP assay.
BCOP	Different Companies Section 6.3.2.2.1	Draize LVET	66 unique 2 unique	30 materials (all from the Draize study) were also tested in the EpiOcular assay

5.1.1 Cytosensor method

The Cytosensor data submitted by the participating companies for anti-microbial cleaning products (and similar formulations) were generated by at least two different protocols. One was the protocol designed for the silicon microphysiometer, the predecessor instrument to the Cytosensor, which is described in Section 2.2.1. This protocol uses a 500 second exposure to cells grown

on a cover slip (see Section 2.2.1 for further explanation). For ease in combining data so that a comprehensive prediction model for both instruments could be developed, data from this protocol were transformed to Cytosensor data by an algorithm described in Section 2.2.1.1.

The second protocol used to generate anti-microbial cleaning products data was the standard Cytosensor protocol used by both the Procter & Gamble Company and the Institute for *In Vitro* Sciences, Inc. This protocol uses an 810 second exposure to cells grown on a Transwell membrane (see Section 2.2.1 for further explanation), and is presented in Annex A1.

Also included in this BRD are data generated from the CTFA Phase III evaluation study on surfactant-based formulations (Gettings, Lordo et al. 1996). This study used the Silicon Microphysiometer protocol (500 sec exposure).

Data from a second large validation study which used surfactants and surfactant-based formulations (some of which were prepared to be identical to the ones used in the CTFA evaluation) – the COLIPA eye irritation study (Brantom, Bruner et al. 1997) - used the Cytosensor protocol (810 sec exposure) which is contained in Annex A2.

5.1.2 EpiOcular method

The EpiOcular data submitted by the participating companies for anti-microbial cleaning products (and similar formulations) were all generated by a single protocol which was developed by the Procter & Gamble Company and Microbiological Associates/IIVS. This protocol uses the EpiOcular tissue model (MatTek Corporation, Ashland, MA) and is contained in Annex A3. See Section 2.2.2.2 for more details on the protocol.

5.1.3 BCOP method

The BCOP data submitted by the participating companies for anti-microbial cleaning products (and similar formulations) were all generated by a common protocol which is contained in Annex A2. This is essentially identical to the "ICCVAM Recommended BCOP Test Method Protocol" which is contained in ICCVAM's test method evaluation report following their review of 4 methods to detect ocular corrosives and severe irritants. The standard exposure time in this protocol is 10 minutes; however, some data are included in this BRD where the corneas were exposed for only three minutes. In fact, it was determined that the 10 minute exposure often overpredicted cleaning formulations which contained >5% solvent. The animal derived toxicity categories were more accurately predicted by a three minute score. Therefore, we suggest that formulations containing >5% solvent be evaluated with a three minute exposure protocol. It is indicated in the text where these types of data are being discussed.

Some interlaboratory variability data are presented in this BRD which were extracted from the ICCVAM BRD on the BCOP assay. The protocols which were used to generate these data are described in the ICCVAM BRD.

5.2 Availability of copies of original data used to evaluate the predictive capacity and reliability of the three test methods

5.2.1 Cytosensor data

For the main analysis of predictive capacity, data from the Cytosensor were submitted by participating companies along with spreadsheets containing the results of animal studies. In some cases, the original reports from the animal studies were submitted. The spreadsheets containing the data are appended to this BRD (Annex C2), and the actual reports can be made available to ICCVAM or the EPA upon request.

For the supplemental information that was used for predictive capacity and reliability (results from the CTFA Phase III evaluation and the COLIPA study), the raw animal data from the CTFA Phase III evaluation are available, but only subsequent transcriptions are available for the COLIPA study. Raw data from the *in vitro* portion of these two studies can be supplied if desired.

5.2.2 EpiOcular data

Raw data for both the *in vitro* and *in vivo* studies reported for the EpiOcular method are available upon the request of ICCVAM or the EPA.

Raw data from the Colgate-Palmolive sponsored validation of the EpiOcular test method (used here for interlaboratory reliability information) can be made available to ICCVAM or the EPA upon request.

5.2.3 BCOP data

Raw data for both the *in vitro* and *in vivo* studies reported for the BCOP method are available upon the request of ICCVAM or the EPA.

Raw data from some of the ancillary studies taken from the ICCVAM BCOP BRD (ICCVAM 2006) may be available from the NICEATM archives.

5.3 Summary of results and prediction models used to evaluate the data

The development of the prediction models for each of the test methods is described in the data analysis section of this BRD (Section 6.0).

5.3.1 Cytosensor test method

Participating companies submitted Cytosensor data for ~275 test samples having formulations similar to those found in typical cleaning product formulations. After evaluating the animal data (all LVET data for these samples) it was found that the data were insufficient to accurately calculate EPA toxicity categories for 170 materials due to termination of the animal test prior to 21 days or individual animal data were not provided. Thus 108 materials remained for which there were both EPA categories and Cytosensor MRD_{50} information. Three of these materials were described as having oxidizing properties and had been tested in the Cytosensor before other studies conducted by the participating companies indicated that oxidizing products often cause a delayed ocular response which is best observed in the BCOP assay. Therefore, the oxidizing materials were not used in the analysis of the Cytosensor performance, leaving 105 unique materials which could be used to gauge the performance of the Cytosensor. Coded information on the 105 materials is given in Table 5-2. Full formulation information on the materials can be traced using the code to identify the appropriate information in Annex B3.

In addition to the company submissions, we were able to obtain Cytosensor and rabbit raw data from 25 materials from the CTFA Phase III eye irritation evaluation study (Gettings, Lordo et al. 1996) on surfactants and surfactant containing materials (Table 5-4). Both LVET and Draize test data were obtained for all 25 materials allowing a comparison of these two rabbit eye test methodologies for deriving the cut-offs needed for a prediction model. A list of the formulations is included in Annex B4.

In order to obtain additional information on the performance characteristics of the CM assay when the traditional Draize test was used to define the EPA and GHS toxicity classification of the formulations, we obtained raw data from a COLIPA-sponsored study (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999) which tested a range of surfactant-containing formulations including 12 surfactants and 7 surfactant-containing materials (Table 5-5). The traditional Draize methodology was used to define the toxicity classifications of the chemicals and formulations.

5.3.1.1 Company Cytosensor data submissions paired with data from the LVET assay

Table 5-2 lists the 105 unique formulations for which both Cytosensor data and rabbit LVET data exist. Table 5-3 summarizes the number of formulations which

fall into each each of the predetermined “buckets”. Both GHS and EPA toxicity categories are listed along with the Cytosensor MRD_{50} value. Where 6-rabbit tests were used, the distribution of 3-rabbit subgroups are listed to indicate the level of variability associated with the final category assignment. See Section 4.8.1 for a discussion of this type of analysis. The protocol used to generate the paired data was the standard Cytosensor protocol used by both the Procter & Gamble Company and the Institute for *In Vitro* Sciences, Inc. This protocol uses an 810 second exposure to cells grown on a Transwell membrane (see Section 2.2.1 for further explanation), and is presented in Annex A1.

Table 5-2 Results of 105 unique materials tested in the Cytosensor assay and the rabbit LVET assay. Four of the materials were tested twice in the LVET assay and have toxicity categories from both tests listed.

Cytosensor Data Paired With LVET-Defined Toxicity Categories																
Code Number	**Physical State**	**Formulation Type**			**Cytosensor MRD_{50}**	***In vivo* GHS**	***In vivo* EPA**	**GHS Categories (3 rabbit subgroups)**				**EPA Categories (3 rabbit subgroups)**			
		#1	**#2**	**#3**	**(mg/mL)**	**(LVET)**	**(LVET)**	**1**	**2A**	**2B**	**NI**	**I**	**II**	**III**	**IV**
1001	liquid	SU	SO		0.435	Non-irritant	Category III	0	0	0	20	0	0	10	10
1002	liquid	SU	SO		0.535	Non-irritant	Category III	0	0	3	17	0	0	20	0
1003	liquid	SU	SO		0.44	Category 2A	Category II	0	16	4	0	0	16	4	0
1004	liquid	SU	SO		0.421	Category 2B	Category III	0	0	1	0	0	0	1	0
1005	liquid	SU	SO		0.411	Category 2A	Category II	0	1	0	0	0	1	0	0
1006	liquid	SU	SO		0.443	Non-irritant	Category III	0	0	0	1	0	0	1	0
1007	liquid	SU	SO		0.428	Category 2B	Category III	0	0	1	0	0	0	1	0
1008	liquid	SU	SO		0.272	Category 2B	Category III	0	0	1	0	0	0	1	0
1009	liquid	SU	SO		0.465	Non-irritant	Category III	0	0	0	20	0	0	19	1
1010	liquid	SU	SO		0.456	Category 1	Category I	1	0	0	0	1	0	0	0
1011	liquid	SU	SO		0.44	Category 1	Category I	1	0	0	0	1	0	0	0
1012	liquid	SU	SO		0.415	Category 2A	Category II	0	1	0	0	0	1	0	0
1013	liquid	SU	SO		0.426	Category 2B	Category III	0	0	1	0	0	0	1	0
1014	liquid	SU			0.444	Non-irritant	Category III	0	0	0	1	0	0	1	0
1015	liquid	SU	SO		0.412	Non-irritant	Category III	0	0	0	1	0	0	1	0
1016	liquid	SU	SO		0.272	Category 2B	Category III	0	0	1	0	0	0	1	0
1017	liquid	SU			0.432	Category 2B	Category III	0	0	1	0	0	0	1	0
1018	liquid	SU			0.465	Category 2B	Category III	0	0	1	0	0	0	1	0
1019	liquid	SU	SO		0.276	Category 1	Category I	1	0	0	0	1	0	0	0
1020	liquid	SU	SO		0.296	Category 1	Category I	1	0	0	0	1	0	0	0
1021	granular	SU	AL		0.19	Non-irritant	Category III	0	0	0	20	0	0	16	4
1022	liquid	SU	SO		0.51	Category 2A	Category I	10	9	1	0	10	9	1	0
1023 (2nd test for 1022)	liquid	SU	SO		0.51	Category 2A	Category III	4	15	1	0	0	0	20	0
1024	liquid	SU	SO		0.2	Category 1	Category I	1	0	0	0	1	0	0	0
1025	liquid	SU	SO		0.829	Non-irritant	Category III	0	0	0	20	0	0	20	0
1026	viscous	SU	SO		0.434	Category 2B	Category III	0	0	20	0	0	0	20	0

Code Number	Physical State	Formulation Type			Cytosensor MRD_{50}	*In vivo* GHS	*In vivo* EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)			
		#1	#2	#3	(mg/mL)	(LVET)	(LVET)	1	2A	2B	NI	I	II	III	IV
1027	liquid	SU	SO		0.44	Non-irritant	Category III	0	0	1	19	0	0	20	0
1028	liquid	SU	SO		0.46	Category 2A	Category III	0	1	0	0	0	0	1	0
1029	liquid	SU	SO		0.45	Category 2B	Category III	0	0	1	0	0	0	1	0
1030	liquid	SU	SO		0.6	Category 2B	Category III	0	0	1	0	0	0	1	0
1031	liquid	SU	SO		0.5	Category 2B	Category III	0	0	1	0	0	0	1	0
1032	liquid	SU	SO		0.96	Category 2A	Category III	0	1	0	0	0	0	1	0
1033 (2nd test for 1032)	liquid	SU	SO		0.96	Category 2B	Category III	0	0	1	0	0	0	1	0
1034	liquid	SU	SO		0.67	Non-irritant	Category III	0	0	0	1	0	0	1	0
1035	liquid	SU	SO		63.9	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1036	liquid	SU	SO		0.79	Non-irritant	Category III	0	0	0	1	0	0	1	0
1037	polymer	SU	AL		9.043	Non-irritant	Category III	0	0	0	1	0	0	1	0
1038	built add	SU	AL		8.916	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1039	liquid	SU	SO		0.26	Category 1	Category I	1	0	0	0	1	0	0	0
1040	liquid	SU	SO		0.76	Category 2A	Category II	0	1	0	0	0	1	0	0
1041	liquid	SU	SO		0.22	Category 2A	Category II	0	1	0	0	0	1	0	0
1042	viscous	SU	SO	AL	22.7	Non-irritant	Category III	0	0	0	1	0	0	1	0
1043	liquid	SU	SO		0.407	Category 2A	Category II	0	20	0	0	0	10	10	0
1044	liquid	SU	SO		0.428	Category 2A	Category II	0	20	0	0	0	19	1	0
1045	liquid	SU	SO		0.344	Category 2A	Category III	0	19	1	0	0	0	20	0
1046	liquid	SU	SO		0.264	Category 2A	Category II	0	1	0	0	0	1	0	0
1047	cream	SU	SO		0.286	Non-irritant	Category III	0	0	0	1	0	0	1	0
1048	liquid	SU	AC		5.81	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1049	liquid	SU	AC		6.02	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1050	liquid	SU	SO	AC	4.99	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1051	liquid	SU	SO		7.103	Category 2B	Category III	0	0	1	0	0	0	1	0
1052	viscous liquid	SU			1.354	Non-irritant	Category III	0	0	0	1	0	0	1	0
1053	liquid	SU			0.0808	Category 2B	Category III	0	0	1	0	0	0	1	0
1054	liquid	SU			0.0773	Category 2B	Category III	0	0	1	0	0	0	1	0
1055	liquid	SU			0.638	Category 2A	Category II	0	1	0	0	0	1	0	0
1056	liquid	SU			0.817	Category 2A	Category II	0	1	0	0	0	1	0	0

Code Number	Physical State	Formulation Type			Cytosensor MRD_{50}	*In vivo* GHS	*In vivo* EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)			
		#1	#2	#3	(mg/mL)	(LVET)	(LVET)	1	2A	2B	NI	I	II	III	IV
1057 (2nd test for 1056)	liquid	SU			0.817	Category 1	Category I	1	0	0	0	1	0	0	0
1058	liquid	SU			0.81	Category 2A	Category II	0	1	0	0	0	1	0	0
1059	liquid	SU			0.787	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1060	liquid	SU			0.9	Non-irritant	Category III	0	0	0	1	0	0	1	0
1061	cream	SU			26.733	Non-irritant	Category III	0	0	0	1	0	0	1	0
1062	cream	SU			46.5	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1063	cream	SU			43.1	Non-irritant	Category III	0	0	0	1	0	0	1	0
1064	liquid	SU			0.501	Non-irritant	Category III	0	0	0	1	0	0	1	0
1065	liquid	SU	SO		300	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1066	liquid	SU			3.8	Non-irritant	Category III	0	0	0	1	0	0	1	0
1067	liquid	SU			2.573	Non-irritant	Category III	0	0	0	1	0	0	1	0
1068	liquid	SU			4.308	Non-irritant	Category III	0	0	0	1	0	0	1	0
1069	liquid	SU			0.556	Non-irritant	Category III	0	0	0	1	0	0	1	0
1070	liquid	SU			1.96	Non-irritant	Category III	0	0	0	1	0	0	1	0
1071	liquid	SU			0.66	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1072	solid/flakes	SU			3.718	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1074	cream	SU			4.19	Non-irritant	Category III	0	0	0	1	0	0	1	0
1075	solid	SU			10.96	Non-irritant	Category IV	0	0	0	20	0	0	0	20
1076	liquid	SU			0.63	Non-irritant	Category III	0	0	4	16	0	0	20	0
1077	liquid	SU			0.63	Category 1	Category I	1	0	0	0	1	0	0	0
1078	gel	SU			0.49	Non-irritant	Category III	0	0	6	14	0	0	20	0
1079	liquid	SU			0.708	Category 2B	Category III	0	0	1	0	0	0	1	0
1080 (2nd test for 1079)	liquid	SU			0.708	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1081	liquid	SU			0.717	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1082	liquid	SU			2.019	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1083	liquid	SU			1.43	Non-irritant	Category III	0	0	0	1	0	0	1	0
1084	liquid	SU			3.86	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1085	liquid	SU			15.18	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1086	liquid	SU			0.93	Category 2B	Category III	0	0	17	3	0	0	20	0
1087	liquid	SU			2.49	Non-irritant	Category IV	0	0	0	1	0	0	0	1

Code Number	Physical State	Formulation Type			Cytosensor MRD_{50}	*In vivo* GHS	*In vivo* EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)			
		#1	#2	#3	(mg/mL)	(LVET)	(LVET)	1	2A	2B	NI	I	II	III	IV
1088	liquid	SO	AL		48.48	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1089	cream	SO			20.652	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1090	liquid	SO			8.085	Non-irritant	Category III	0	0	0	1	0	0	1	0
1091	liquid	AC	SU		6.41	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1092	liquid	SO	SU		300	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1093	liquid	SO			5.97	Category 2B	Category III	0	0	1	0	0	0	1	0
1094	liquid	SO			142.857	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1095	liquid	SO			69.842	Non-irritant	Category III	0	0	0	1	0	0	1	0
1096	liquid	SO			22.438	Non-irritant	Category III	0	0	0	1	0	0	1	0
1097	liquid	SO			22.172	Non-irritant	Category III	0	0	0	1	0	0	1	0
1098	creamy liquid	SO			20.68	Non-irritant	Category IV	0	0	0	20	0	0	0	20
1099	creamy liquid	SO	SU		3.96	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1100	liquid	SO			18.834	Non-irritant	Category III	0	0	0	1	0	0	1	0
1101	liquid	SO			16.581	Non-irritant	Category III	0	0	0	1	0	0	1	0
1102	liquid	SO	SU		0.92	Category 2B	Category III	0	0	1	0	0	0	1	0
1103	semi-viscous liquid	SO	SU		21.9	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1104	liquid	AL	SO		41.5	Non-irritant	Category III	0	0	0	1	0	0	1	0
1105	liquid	AL	SO		69.63	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1106	liquid	AL	SO		52.13	Non-irritant	Category III	0	0	0	1	0	0	1	0
1107	liquid	AL	SO		21.4	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1108	viscous liquid	SO	SU	AC	2.2	Non-irritant	Category IV	0	0	0	1	0	0	0	1
1109	thin liquid	SO	AC	SU	3.377	Non-irritant	Category III	0	0	0	1	0	0	1	0
1110	liquid	SO	AC		30.365	Non-irritant	Category III	0	0	0	1	0	0	1	0

AC=Acid, AL=Alkaline (base), SO=Solvent, SU=Surfactant

Table 5-3 gives the distribution of materials in Table 5-2. It is obvious that the distribution of product categories is relatively uneven, but follows a pattern similar to that of the types of anti-microbial cleaning products on the market (personal communication, P&G).

Table 5-3 Distribution of product categories originally submitted with both animal eye irritation data and Cytosensor data.

Product Categories	Number of products tested
Surfactants	82
Acids	1
Bases	4
Solvents	18
Total	105

5.3.1.2 CTFA Phase III study (Gettings, Lordo et al. 1996)

The CTFA Phase III study was chosen for inclusion in this BRD since it is helpful to understand how the Draize and the LVET perform on a set of materials (surfactant-based personal care products) for which there are CM data and which are similar to those materials contained in this BRD. The animal data can be found in Annexes C3-C6.

The CTFA Phase III study (Gettings, Lordo et al. 1996) was an evaluation program of a number of *in vitro* eye irritation tests. The project's original goal was to determine how well the *in vitro* tests predicted the Draize MAS scores for 25 surfactant-based personal care products, but a secondary analysis conducted at the conclusion of the primary study included LVET MAS scores as well. The reference data for the CTFA Phase III study are arguably the most useful of the animal data from any of the studies in this BRD. Data from both the Draize and LVET assays were obtained under GLP-compliant conditions and with a randomized block design utilizing three male and three female rabbits for each chemical. There are several advantages to the block design: 1) it simulates to some extent within lab day-to-day variability since for each chemical not all rabbits are dosed on the same day, and 2) it eliminates some of the scoring bias since the scorers read each animal independently and are unaware of which six rabbits were treated with the same test article. However, the main positive point about the study is that the *in vitro* and *in vivo* assays were run nearly concurrently (separated only by a few weeks) using samples from the same batch of chemical or formulation. The one negative point to this study is that ocular anesthesia was used during the rabbit test (both Draize and LVET) and to the best of our knowledge none of the other animal assays in this BRD used ocular anesthesia. There are reports that rabbits given ocular anesthesia may have a more intense ocular reaction than animals treated without anesthesia, *e.g.,* Gunderson & Liebmann (1944).

Only one laboratory (Microbiological Associates, Inc., Rockville, MD) contributed CM data for this study. All 25 chemicals in the study were deemed compatible for testing with the CM. An overall summary of the CTFA Phase III study including the chemical identities, animal scores, and *in vitro* scores is given in Table 5-4. Although these studies were conducted with the silicon microphysiometer, for ease of comparison with the other studies in this section of the BRD, the *in vitro* MRD_{50} values have been converted to CM values using the relationship presented in Section 2.2.1.1.

Table 5-4 shows that in the CTFA Phase III study most materials (16/25; 64%) are assigned the same EPA toxicity category by either the LVET or the Draize test, supporting the fact that Draize and LVET are not all that different. The total concordance is 64%, with 12% differing by one category and 24% differing by 2 categories.

A similar analysis by GHS categories shows that there is 64% concordance, with 4% differing by one category, 16% differing by 2 categories and 16% differing by three categories.

Table 5-4 Summary of Cytosensor data from the CTFA Phase III study using toxicity classifications determined by both the Draize Rabbit Test and the Low Volume Eye Test for surfactant-containing materials (Gettings, Lordo et al. 1996)

CTFA Phase III Cytosensor In Vitro Data DRAIZE & LVET *In Vivo* Eye Classifications									
CTFA chemical number	**Substance**	**Test Code**	**Concentration Tested**	***In Vivo* GHS[1,2] (DRAIZE)**	***In Vivo* GHS[1,2] (LVET)**	***In Vivo* EPA[3,4] (DRAIZE)**	***In Vivo* EPA[3,4] (LVET)**	**DRAIZE[9] MMAS**	**CM converted value MRD_{50} (mg/mL)**
1	Shampoo 7	HZA	100%	Category 1	No category	Category I	Category III	37.8	1.18
2	Liquid Soap 1	HZB	25%	No category	No category	Category III	Category IV	20.7	2.80
3	Shampoo 1	HZC	25%	No category	No category	Category III	Category III	36.0	1.72
4	Shampoo 5	HZD	25%	No category	No category	Category III	Category III	19.5	2.78
5	Gel Cleanser	HZE	100%	No category	No category	Category I	Category III	22	3.19
6	Baby Shampoo 2	HZF	100%	Category 1	No category	Category I	Category III	37.5	1.50
7	Shampoo 8	HZG	25%	No category	No category	Category III	Category III	17.8	2.80
8	Eye Makeup re.	HZH	100%	No category	No category	Category IV	Category IV	2.3	20.0
9	Skin Cleaner	HZI	100%	Category 1	Category 2B	Category I	Category I	41.0	1.09
10	Mild Shampoo	HZJ	100%	No category	No category	Category IV	Category IV	8.2	6.38
11	Bubble bath	HZK	100%	Category 1	Category 2B	Category I	Category I	39.7	0.97
12	Foam Bath	HZL	100%	Category 1	No category	Category I	Category III	37.8	1.09
13	Shampoo 3	HZM	25%	No category	No category	Category III	Category III	12.7	3.11
14	Shampoo 6	HZN	25%	No category	No category	Category III	Category III	18.0	2.56
15	Baby Shampoo 1	HZP	100%	No category	No category	Category III	Category III	11.7	2.45
16	Cleansing Gel	HZQ	100%	No category	No category	Category III	Category IV	17.2	5.85
17	Facial Cleansing Foa	HZR	25%	No category	No category	Category I	Category III	39.0	5.60
18	Shower Gel	HZS	100%	Category 1	Category 2B	Category I	Category I	41.4	1.13
19	Polishing Scrub	HZT	100%	No category	No category	Category IV	Category IV	7.0	30.9
20	Hand Soap	HZU	25%	No category	No category	Category III	Category III	33.7	4.85
21	Shampoo 4	HZV	25%	No category	No category	Category III	Category III	25.2	2.34
22	Liquid Soap 2	HZW	25%	2B	No category	Category III	Category III	31.0	2.64
23	Shampoo 2	HZX	100%	Category 1	No category	Category I	Category III	40.0	1.20
24	Shampoo AntiD	HZY	100%	Category 1	Category 2B	Category I	Category II	43.0	1.14
25	Facial Cleanser	HZZ	100%	No category	No category	Category IV	Category IV	3.7	>168.9

[1]GHS=Globally Harmonized System (UN [2003])

[2]Eye Irritant Category 1 = irreversible effects on the eye/serious damage to the eye; Category 2A = reversible effects on the eye/irritating to the eyes; Category 2B = reversible effects on the eye/mildly irritating to the eyes; No category = no effects on the eye

[3]EPA=U.S. Environmental Protection Agency (EPA [1996]).

[4]Toxicity Category I for the Primary Eye Irritation Study = Corrosive, or corneal involvement or irritation not reversible within 21 days; Category II = Corneal involvement or irritation clearing in 8-21 days; Category III = Corneal involvement or irritation clearing in 1-7 days; Category IV: minimal effects clearing in less than 24 hr

[5]MMAS scores reported in Gettings et al. (1996)

5.3.1.3 COLIPA Validation study for eye irritation

In 1995/1996 the European Cosmetics, Toiletry and Perfumery Association (COLIPA) sponsored an international validation study of *in vitro* eye irritation methods (Brantom, Bruner et al. 1997). The COLIPA study used a set of 55 cosmetic formulations and ingredients - a large proportion of which were pure surfactants or surfactant based formulations - to assess the ability of *in vitro* methods to predict eye irritation potential. Two laboratories conducted the CM assay according to a standardized protocol (Annex A2) which used an 810 second exposure time. Raw data from the studies conducted by Microbiological Associates, Inc. and CellTox AB were obtained from the archives of the Institute for *In Vitro* Sciences, Inc. Mean data from these two laboratories for each chemical are presented in Table 5-5.

The reference data for the COLIPA study came from three main sources; two for the neat chemicals and one for the formulations. The data for the chemicals came from the ECETOC data bank (ECETOC 1992) and the EU isolated cornea study (Gautheron, Giroux et al. 1994). All of these data are now available in a new edition of the ECETOC data bank (ECETOC 1998). The raw animal data are also found in Annexes C7&C8.

Thirty-two formulations were used in the COLIPA study, and the Draize scores for these formulations come from Draize tests conducted contemporaneously with this study. The formulations were newly prepared for the COLIPA study, but most were based on formulations that had been tested in Phases I, II, and III of the CTFA evaluation program (Feder, Lordo et al. 1991; Gettings, Dipasquale et al. 1994; Gettings, Lordo et al. 1996). Thus, it is likely that for the formulations, the *in vitro* tests were challenged with exactly the same material as the *in vivo* test. The same cannot be said for the chemicals since historical data were used for them. Because the evaluation of formulations (anti-microbial cleaning products) is the focus of this BRD, only the results with the formulations, or with pure surfactants, from the COLIPA study will be addressed here.

There were 19 surfactants and surfactant-containing materials which had data from the two participating CM laboratories. An overall summary of the COLIPA study including the chemical identities, animal scores and *in vitro* scores (averages from MA and CellTox AB) are given in Table 5-5. The formulations are included in Annex B5.

Table 5-5 Summary of Cytosensor and *in vivo* data from the COLIPA study which includes average values (see footnotes) from MA and CellTox AB laboratories (Brantom, Bruner et al. 1997).

COLIPA study - Surfactants and Sufactant-based Formulations *Cytosensor and In Vivo* Eye Irritation Classifications							
COLIPA chemical number	**Substance**	**Concentration Tested**	**n. of animals**	***In Vivo* GHS[1,2]**	***In Vivo* EPA[3,4]**	**ECETOC MMAS Score[5]**	**Average MRD_{50} (mg/mL)**
5	Shampoo no. 1 - normal	100%	3	Category 1	Category I	33.3	0.735
6	Eye make-up remover	100%	3	No Category	Category IV	0.7	93.5
11	Polyethylene glycol 400	100%	6	No Category	Category IV	0.0	306.4
13	Triton X-100	1%	3	No Category	Category III	1.7	19.0
15	Tween 20	100%	4	No Category	Category III	4.0	6.50
17	Sodium lauryl sulphate	3%	6	No Category	Category III	16.0	3.00
20	Triton X-100 [2]	5%	6	Category 2A	Category III	32.3	3.54
21	Benzalkonium chloride [1]	1%	4	Category 2A	Category I	34.3	4.22
21	Benzalkonium chloride [2]	1%	6	Category 1	Category I	56.3	4.22
23	Sodium lauryl sulphate	15%	6	Category 1	Category I	59.2	0.513
24	Sodium lauryl sulphate	30%	6	Category 2A	Category II	60.5	0.312*
25	Triton X-100	10%	6	Category 1	Category II	59.0	1.85
26	Benzalkonium chloride	5%	4	Category 1	Category I	83.8	1.095
27	Benzalkonium chloride	10%	3	Category 1	Category I	108.0	0.314
28	Pump deodorant / antiperspirant	100%	3	No Category	Category III	14.7	33.54
34	Gel cleanser	100%	3	No Category	Category III	15.7	5.58
36	Shampoo - baby	100%	3	Category 1	Category I	36.0	2.33
39	Liquid soap no.1	100%	3	Category 1	Category I	37.0	0.78
49	Skin cleanser	100%	3	Category 1	Category I	34.3	0.70
52	Cetylpyridinium bromide	6%	4	Category I	Category I	85.8	1.36*

* - MA value only, CellTox AB designated unsuitable for testing

[1]GHS=Globally Harmonized System (UN [2003])

[2]Eye Irritant Category 1 = irreversible effects on the eye/serious damage to the eye; Category 2A = reversible effects on the eye/irritating to the eyes; Category 2B = reversible effects on the eye/mildly irritating to the eyes; No category

[3]EPA=U.S. Environmental Protection Agency (EPA [1996])

[4]Toxicity Category I for the Primary Eye Irritation Study = Corrosive, or corneal involvement or irritation not reversible within 21 days;

[5]MMAS scores reported in Harbell et al. (1999)

5.3.2 EpiOcular

Participating companies submitted EpiOcular data for 61 test samples having formulations similar to those found in typical cleaning product formulations. The raw animal data can be found in Annex C1. After evaluating the animal data (both LVET data and Draize data) it was found that the animal data were insufficient to accurately calculate EPA toxicity Categories for 6 materials due to termination of the animal test prior to 21 days or individual animal data were not provided. Thus 55 materials remained for which there were both EPA Categories and EpiOcular ET_{50} information. Twenty-five materials were paired with LVET data (Table 5-8) and 30 were paired with Draize data (Table 5-6). Tables 5-7 and 5-9 give the distribution of materials in Tables 5-6 and 5-8, respectively.

Data from another set of studies conducted to validate the EpiOcular assay were also submitted for this BRD. Seventy-three surfactants or surfactant-based materials (or dilutions of materials) were tested in these studies. However, the EpiOcular protocol used in those studies differs (a dilution of the test material was performed before the testing) from the protocol being proposed in this BRD; therefore, these studies will be presented only as supporting information for interlaboratory reproducibility (Section 7.2.3).

Table 5-6 EpiOcular data paired with the Draize test

EpiOcular Data Paired With DRAIZE - Defined Toxicity Categories															
Code Number	*In Vivo* Dosing Volume	Formulation Type			In Vivo GHS	In Vivo EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)				EpiOcular ET_{50} (min)
		#1	#2	#3	(DRAIZE)	(DRAIZE)	1	2A	2B	NI	I	II	III	IV	
H	0.1	AL	SU		Non-irritant	Category II	0	9	1	10	0	10	10	0	9.4
I	0.1	SU	AL		Non-irritant	Category III	0	0	0	20	0	0	10	10	12
J	0.1	SU			Non-irritant	Category III	0	0	0	20	0	0	20	0	19.3
K	0.1	RC	SU		Non-irritant	Category IV	0	0	0	20	0	0	0	20	> 240
P	0.1	Phenolic	AL		Non-irritant	Category IV	0	0	0	1	0	0	0	1	125.8
R	0.1	SU			Non-irritant	Category IV	0	0	0	20	0	0	0	20	> 240
T	0.1	AC			Non-irritant	Category IV	0	0	0	1	0	0	0	1	31.6
W	0.1	SU			Non-irritant	Category IV	0	0	0	20	0	0	0	20	39.6
CJ	84 mg solid				Category 1	Category I	1	0	0	0	1	0	0	0	2.9
AG	0.1	AL			Category 1	Category I	20	0	0	0	20	0	0	0	<0.17
AH	0.1	AL	SU		Category 1	Category I	18	2	0	0	19	0	1	0	0.4
AI	0.03	AL	SU		Category 1	Category I	16	4	0	0	16	0	4	0	<0.17
AJ	0.03	AL	SU		Category 1	Category I	20	0	0	0	20	0	0	0	<0.17
AK	0.1	AL	SO	SU	Category 1	Category I	20	0	0	0	20	0	0	0	<0.17
AL	0.03	AL	SO	SU	Category 2A	Category I	10	10	0	0	10	0	10	0	<0.17
AM	0.1	SO	AL		Category 1	Category I	20	0	0	0	20	0	0	0	<0.17
AN	0.03	AL	SU		Category 1	Category I	19	1	0	0	16	4	0	0	1.5
AO	0.03	AL	SO	SU	Category 1	Category I	20	0	0	0	20	0	0	0	<0.17
AP	0.03	AL	SU		Category 1	Category I	16	4	0	0	16	0	4	0	<0.17
AT	0.1	RC	AL		Category 1	Category I	20	0	0	0	19	1	0	0	<1
AU	0.1	RC	AL		Category 1	Category I	20	0	0	0	20	0	0	0	<1
AV	0.1	RC	AL		Category 1	Category I	1	0	0	0	1	0	0	0	<1
AX	0.03	SO	AL		Category 1	Category I	19	1	0	0	16	3	1	0	<0.17
BB	0.1	SO			SCNM	Category IV	0	0	0	0	0	0	0	20	>240
BE	0.1	AC	SU		Non-irritant	Category III	9	0	0	11	0	0	16	4	4
BJ	0.1	AL	SU		Non-irritant	Category III	0	0	10	10	0	0	20	0	2.1
BK	0.1	SO			Non-irritant	Category III	0	0	0	1	0	0	1	0	9.4
BM	0.1	SO			Non-irritant	Category IV	0	0	0	20	0	0	0	20	4.9
BL	0.1	SO			Non-irritant	Category IV	0	0	0	20	0	0	0	20	6.7
BN	0.1	SU			Non-irritant	Category IV	0	0	0	1	0	0	0	1	1.8

AC = Acid; AL = Alkaline (base); RC = Reactive Chemistry (Oxidizer); SO = Solvent; SU = Surfactant; SCNM = Study Criteria Not Met

Table 5-7 Distribution of product categories for EpiOcular data paired with the Draize test

Product Categories	Number of products tested
Surfactants	5
Acids	2
Alkaline	11
Oxidizers	4
Solvent	6
Other	2
Total	30

Table 5-8 EpiOcular data paired with LVET data

EpiOcular Data Paired With LVET - Defined Toxicity Categories															
Code Number	Formulation Type			*In Vivo* GHS	*In Vivo* EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)				EpiOcular ET_{50} (min)	LVET MAS
	#1	#2	#3	(LVET)	(LVET)	1	2A	2B	NI	I	II	III	IV		
CY	SU	SO		Category 1	Category I	1	0	0	0	1	0	0	0	2.85	
DC	RC	SU		Category 1	Category I	1	0	0	0	1	0	0	0	1.1	59.67
DH	RC	SU		Category 1	Category I	1	0	0	0	1	0	0	0	0.7	60
DD	RC	SU		Category 2A	Category II	0	1	0	0	0	1	0	0	0.9	49.333
CK	SU			Non-irritant	Category III	0	0	0	1	0	0	1	0	21.75	6
CN	SO			Category 2B	Category III	0	0	1	0	0	0	1	0	49.5	18.333
CQ	SU			Non-irritant	Category III	0	0	0	1	0	0	1	0	29.5	13
CS	SU			Non-irritant	Category III	0	0	0	1	0	0	1	0	23.8	4
CU*	SU	AL		Non-irritant	Category III	0	0	0	20	0	0	16	4	20.25	5.5
CV	RC	AL	SU	Category 2A	Category III	0	1	0	0	0	0	1	0	16	11.5
CW*	SU	SO		Non-irritant	Category III	0	0	0	20	0	0	20	0	13.7	10.2
CX	SU	SO		Non-irritant	Category III	0	0	0	1	0	0	1	0	11.2	21.7
DB	RC	AL		Non-irritant	Category III	0	0	0	1	0	0	1	0	1.7	7
DG*	SU	SO		Category 2B	Category III	0	0	20	0	0	0	20	0	0.75	27.2
DI*	SU			Non-irritant	Category III	0	0	4	16	0	0	20	0	0.484	17.2
DK	RC	AL		Category 2A	Category III	0	1	0	0	0	0	1	0	0.167	33
CO	SO			Non-irritant	Category IV	0	0	0	1	0	0	0	1	47.6	4
CP	SU	SO		Non-irritant	Category IV	0	0	0	1	0	0	0	1	29.5	0
CR	SU	SO		Non-irritant	Category IV	0	0	0	1	0	0	0	1	26.1	2.667
CT	SU	SO		Non-irritant	Category IV	0	0	0	1	0	0	0	1	20.8	2.667
CZ	RC	SO	AC	Non-irritant	Category IV				1				1	2.1	0
DA	RC	SU	SO	Non-irritant	Category IV				1				1	1.9	0
DE	RC	SO		Non-irritant	Category IV				1				1	0.85	0
DF	SO			Non-irritant	Category IV	0	0	0	1	0	0	0	1	0.8	1.333
DJ	SO			Non-irritant	Category IV	0	0	0	1	0	0	0	1	0.45	1.333

* Six animal subgroups were used to determine GHS and EPA categories
AC = Acid; AL = Alkaline (base); RC = Reactive Chemistry (Oxidizer); SO = Solvent; SU = Surfactant

Table 5-9 Distribution of product categories for EpiOcular data paired with the LVET test

Product Categories	Number of products tested
Surfactants	12
Acids	0
Alkaline	0
Oxidizers	9
Solvent	4
Other	0
Total	25

5.3.3 BCOP

5.3.3.1 Data from participating companies

Participating companies submitted BCOP data for 38 test samples having formulations similar to those found in typical cleaning product formulations. The raw animal data can be found in Annex C1. After evaluating the animal data (all Draize data for these samples), it was found that they were insufficient to accurately calculate EPA toxicity Categories for 8 materials due to termination of the animal test prior to 21 days or individual animal data were not provided. Thus 30 materials remained for which there were both EPA Categories and BCOP information. These 30 materials are highlighted in Table 5-10.

In addition to the company submissions, we were able to obtain raw data from 25 materials from the CTFA Phase III study (Gettings, Lordo et al. 1996) on surfactants and surfactant containing materials (which are similar to the materials used in many anti-microbial cleaning products). Both LVET and Draize test data were obtained for all 25 materials allowing a comparison between these two rabbit eye test methodologies.

We were also able to obtain raw data from the European Commission/British Home Office (EC/HO) study (Balls, Botham et al. 1995) which tested a range of materials including 15 surfactants. All animal studies (historically derived data) were conducted with the traditional Draize methodology. Table 5-10 details the BCOP data from participating companies paired with Draize-defined toxicity categories. Table 5-11 gives the distribution of the BCOP data from Table 5-10.

Table 5-10 BCOP data from participating companies paired with Draize-defined toxicity categories (with the exception of two materials which were defined using the LVET assay). Highlighted materials were the original 30 materials submitted.

BCOP Data Paired With DRAIZE - Defined Toxicity Categories																	
Code Number	In Vivo Dosing Volume	Formulation Type			In Vivo GHS	In Vivo EPA	GHS Categories (3 rabbit subgroups)				EPA Categories (3 rabbit subgroups)				BCOP IV Score	
		#1	#2	#3	(DRAIZE)	(DRAIZE)	1	2A	2B	NI	I	II	III	IV	10 min	3 min
A	0.1	SU	AL		Category 1	Category I	16	4	0	0	16	4	0	0	206.9	132.8
B	0.1	SU	AL		Category 1	Category I	20	0	0	0	20	0	0	0	152.2	108
C	0.1	RC	SU	AC	Category 1	Category I	16	0	0	4	16	0	4	0	29.7	10.3
D	0.1	AC			Category 1	Category I	1	0	0	0	1	0	0	0	187.7	67.5
E	0.1	SU	AL		Category 1	Category I	1	0	0	0	1	0	0	0	196.2	110.5
F	0.1	RC	SU	AC	Category 1	Category I	1	0	0	0	1	0	0	0	360.8	18.2
G	0.1	SU/SO			Category 1	Category I	20	0	0	0	20	0	0	0	139.7	133.3
H	0.1	AL	SU		Non-irritant	Category II	0	9	1	10	0	10	10	0	14	2.85
I	0.1	SU	AL		Non-irritant	Category III	0	0	0	20	0	0	10	10	0.6	-0.3
J	0.1	SU			Non-irritant	Category III	0	0	0	20	0	0	20	0	7.7	2.6
K	0.1	RC	SU		Non-irritant	Category IV	0	0	0	20	0	0	0	20	0.3	0
L	0.1	SU			Non-irritant	Category III	0	0	0	20	0	0	19	1	5.5	2.8
M	0.1	SU	AL		Non-irritant	Category III	0	0	10	10	0	0	20	0	55.7	15.3
N	0.1	RC			Non-irritant	Category III	0	0	0	1	0	0	1	0	152.7	7.2
O	0.1	SU	AL		Non-irritant	Category IV	0	0	0	20	0	0	0	20	7.2	2.6
P	0.1	Phenolic	AL		Non-irritant	Category IV	0	0	0	1	0	0	0	1	1.1	-0.3
Q	0.1	SU			Non-irritant	Category IV	0	0	0	1	0	0	0	1	13.5	3.3
R	0.1	SU			Non-irritant	Category IV	0	0	0	20	0	0	0	20	0.2	-0.6
S	0.1	AC	SU		Non-irritant	Category IV	0	0	0	1	0	0	0	1	18.8	9.2
T	0.1	AC			Non-irritant	Category IV	0	0	0	1	0	0	0	1	1.8	0
U	0.1	SU	AL		Non-irritant	Category IV	0	0	0	20	0	0	0	20	3.4	2.9
V	0.1	SU/SO			Non-irritant	Category IV	0	0	0	20	0	0	0	20	20.8	3.5
W	0.1	SU			Non-irritant	Category IV	0	0	0	20	0	0	0	20	5.7	3.5
X	0.1	RC	AL		Category 2A	Category I	9	8	0	3	10	0	10	0	81.9	41.7
Y	0.1	RC	AL		Category 2A	Category II	0	19	1	0	0	16	4	0	74.9	65
Z	0.1	SO			Category 2A	Category II	0	1	0	0	0	1	0	0	31.6	12.8
AB	0.1	SU			Category 1	Category I	20	0	0	0	20	0	0	0	90	51.9
AC	0.1	AC			Category 1	Category I	15	0	1	4	16	0	4	0	134.8	101
AD	0.1	SU			Category 1	Category I	20	0	0	0	20	0	0	0	113.1	79
AE	0.1	AL			Category 1	Category I	1	0	0	0	1	0	0	0	66.7	37.8
AF	0.1	AC			Non-irritant	Category IV	0	0	0	1	0	0	0	1	9.6	2.5
AG	0.1	AL			Category 1	Category I	20	0	0	0	20	0	0	0	391.9	
AH	0.1	AL	SU		Category 1	Category I	18	2	0	0	19	0	1	0	255.7	
AI	0.03	AL	SU		Category 1	Category I	16	4	0	0	16	0	4	0	354.7	
AJ	0.03	AL	SU		Category 1	Category I	20	0	0	0	20	0	0	0	357.1	
AK	0.1	AL	SO	SU	Category 1	Category I	20	0	0	0	20	0	0	0	444.3	
AL	0.03	AL	SO	SU	Category 2A	Category I	10	10	0	0	10	0	10	0	353.6	
AM	0.1	SO	AL		Category 1	Category I	20	0	0	0	20	0	0	0	135.8	
AN	0.03	AL	SU		Category 1	Category I	19	1	0	0	16	4	0	0	113.5	
AO	0.03	AL	SO	SU	Category 1	Category I	20	0	0	0	20	0	0	0	216.2	
AP	0.03	AL	SU		Category 1	Category I	16	4	0	0	16	0	4	0	393.3	
AQ	0.1	RC	AL	SU	Category 1	Category I	20	0	0	0	20	0	0	0	84.9	47.8
AR	0.1	RC	AL	SU	Category 1	Category I	20	0	0	0	20	0	0	0	116.1	72.1
AS	0.1	RC	AL	SU	Category 1	Category I	20	0	0	0	19	1	0	0	79.8	36.6
AT	0.1	RC	AL		Category 1	Category I	20	0	0	0	19	1	0	0	85.6	49.8
AU	0.1	RC	AL		Category 1	Category I	20	0	0	0	20	0	0	0	122.2	64.5
AV	0.1	RC	AL		Category 1	Category I	1	0	0	0	1	0	0	0	191.8	68.8
AW	0.1	RC	AL		Category 1	Category I	19	1	0	0	19	1	0	0	43.1	29.6
AX	0.03	SO	AL		Category 1	Category I	19	1	0	0	16	3	1	0	157.3	
AY	0.1	RC	AL		Category 1	Category I	1	0	0	0	1	0	0	0	194.3	79.7
BB	0.1	SO			SCNM	Category IV	0	0	0	0	0	0	0	20	2	0
BD	0.1	SO	SU	AL	Non-irritant	Category III	0	0	0	20	0	0	20	0	18.3	2.6
BE	0.1	AC	SU		Non-irritant	Category III	9	0	0	11	0	0	16	4	15	
BF	0.1	SO	AC	SU	Category 2A	Category III	0	1	0	0	0	0	1	0	63.5	30.4
BJ	0.1	AL	SU		Non-irritant	Category III	0	0	10	10	0	0	20	0	78.3	
BK	0.1	SO			Non-irritant	Category III	0	0	0	1	0	0	1	0	6.7	2.6
BL	0.1	SO			Non-irritant	Category IV	0	0	0	20	0	0	0	20	6	7.7
BM	0.1	SO			Non-irritant	Category IV	0	0	0	20	0	0	0	20	25.4	11.6
BN	0.1	SU			Non-irritant	Category IV	0	0	0	1	0	0	0	1	13.5	
BP	0.1	SO			Non-irritant	Category IV	0	0	0	1	0	0	0	1	19.1	3.9
BQ	0.1	SO			Non-irritant	Category IV	0	0	0	1	0	0	0	1	33.6	16.8
BR	0.01*	SU	SO		Non-irritant	Category IV	0	0	0	20	0	0	0	20	23.2	4.5
BS	0.01*	RC	AL		Category 2A	Category III	0	1	0	0	0	0	1	0	273.6	121.3
CG	aerosol	AL	SU		Non-irritant	Category IV	0	0	0	1	0	0	0	1	3.9	3.5
CH	0.1	SO			Non-irritant	Category III	0	0	0	1	0	0	1	0	17.4	
EF		RC			Category 2A	Category II	0	20	0	0	0	10	10	0	104.8	54.5
EG	0.1	AC			Category 2A	Category II	0	1	0	0	0	1	0	0	71.8	57.5
BJ	0.1	AL	SU		Non-irritant	Category III	0	0	10	10	0	0	20	0	54.6	8.2

* = Materials tested in the LVET assay

Table 5-11 Distribution of materials conducted in the BCOP assay.

Product Categories	Number of products tested
Surfactants	18
Acids	7
Alkaline	14
Oxidizers	16
Solvent	12
Other	1
Total	68

5.4 Use of coded chemicals and compliance with GLP Guidelines

5.4.1 Company-submitted anti-microbial cleaning product *in vitro* data

Because some of the *in vitro* data were submitted to IIVS in spreadsheets, it was impossible to determine which data were generated under GLP compliance and which were not. However, all of the BCOP data (Section 6.3.2.2.1) generated after the original submissions were conducted with full GLP compliance.

Essentially all of the company-submitted *in vitro* data generated for anti-microbial cleaning products and similar formulations were generated using coded chemicals.

5.4.2 Data obtained from secondary sources

Both *in vitro* and *in vivo* data obtained from publications or internal records for the CTFA Phase III study (Gettings, Lordo et al. 1996) were generated with full GLP compliance. Coded test materials were used for both the *in vitro* and *in vivo* portion of this study.

In vitro data from the COLIPA study (Brantom, Bruner et al. 1997) were generated with full GLP compliance, but some of the *in vivo* data were obtained from historical sources so it could not be determined whether or not all of these tests were done with GLP compliance. The data for formulations conducted in the COLIPA study were generated with coded test materials, but it could not be determined if all of the substances were tested as coded materials.

6 Test Method Predictive Capacity

Prediction models for each of the three *in vitro* assays were constructed using the same approach (a graphical one). For each model all the paired *in vitro* and *in vivo* data provided were used, and the *in vitro* data were plotted against the *in vivo*-defined toxicity category (both EPA and GHS). In some cases only LVET data was available and in other cases only Draize data. Generally each type of data was analyzed separately, although we generally concluded that the prediction models were the same regardless of the *in vivo* assay used.

Once the data were graphed, cut-off lines were fitted by eye to provide the "best" predictions. A description of these cut-offs then became the prediction model. Our strategy in setting the cut-offs was to minimize under predictions of toxicity at the expense of over predictions. Of course, over and under predictions are somewhat arbitrary terms since we have shown earlier in this BRD (Section 4.8.1) that repeated three-rabbit eye irritation tests do not necessarily provide identical toxicity classifications. In other words, a second rabbit test may over or under predict the first test.

Although data from the testing of anti-microbial cleaning products (and related cleaning products) were primarily used to set the cut-offs, additional data from chemically related formulations and some pure substances (e.g. surfactants) were used to provide supporting information for our decisions.

6.1 Cytosensor predictive capacity

6.1.1 Using the LVET assay to define a prediction model for the CM

<u>EPA Labeling Categories</u>

The distribution of product categories originally submitted with both animal eye irritation data (LVET) and Cytosensor *in vitro* data is shown in Table 6-1. It can be seen that there were significantly more surfactants than any other product category tested with the Cytosensor. No oxidizing formulations were tested using the Cytosensor.

Table 6-1 Distribution of product categories originally submitted with both animal eye irritation data and Cytosensor *in vitro* data.

Product Categories	Number of products tested
Oxidizers	0
Surfactants	82
Acids	1
Bases	4
Solvents	18
Total	105

As the first step towards determining a prediction model for CM data, we created a scatter plot showing the MRD_{50} for each material plotted against the EPA labeling category that had been determined by an LVET assay. Figure 6-1 shows the distribution of MRD_{50} values for all of the 105 antimicrobial cleaning products for which paired animal data and CM data were available). We then wished to determine if the results were distributed in such a way that the products with different EPA labeling categories could be easily separated. It is immediately apparent from Figure 6-1 that the distribution of MRD_{50} scores across the EPA labeling categories is not random. EPA Category I and II materials all have MRD_{50}'s <1 mg/mL., and only Category IV materials have MRD_{50} values >80 mg/mL. This distribution allowed us to set cut-off values by eye for predicting EPA labeling categories. We attempted to choose cut-offs conservatively with a bias towards having as few under predictions as was reasonable. No statistical methods were employed to construct the proposed prediction model.

Beginning with predictions of the most severe labeling categories, we found that it was not possible to envision a cut-off value that would distinguish Category I materials from Category II materials because of the significant overlap of their MRD_{50} values. Thus we chose to identify all materials in both of the highest toxicity categories with a conservatively set cut-off value of 2.0 mg/mL. A materials whose MRD_{50} value is <2.0 mg/mL will be labeled as an EPA I. MRD_{50} values of all Category I & II materials in this dataset fall below this cut-off. Thus all materials with MRD_{50} values below 2.0 mg/mL must be given the most severe designation – Category I.

Similarly, MRD_{50} values for EPA Category III and IV materials have significant overlap, although at least three of the Category IV materials have MRD_{50} scores ≥ 80 mg/mL. Thus it is possible to suggest an upper cut-off limit of MRD_{50} > 80 mg/mL to separate some EPA Category IV materials from Category III materials. Materials whose MRD_{50} values are ≥2 mg/mL and <80 mg/mL are defined as being EPA Category III. No animal-defined Category I or II materials are underpredicted by this proposed prediction model. However, since many Category III materials and a few of the Category IV materials fall below the 2.0 mg/mL proposed cut-off for Category I materials, many EPA Category III and a few EPA Category IV materials will be over predicted, and hence over labeled. This outcome has been accepted by the manufacturers who have co-authored this BRD.

Figure 6-1 shows a plot of MRD_{50} values versus EPA category assignments (by LVET) with the above-proposed cut-off values added. Included in Figure 6-1 are three materials for which two sets of animal results were available. Data from both LVET trials have been included to underscore the variability of the animal test and indicate that no *in vitro* test can be expected to predict a given animal score any better than a second animal test itself might be expected to do. It can be seen that for Material 1022, for example, the results of the two animal tests differed by two full classifications (an EPA I versus an EPA III)! The two other materials each differed

by a single category. Formulation 1056 was categorized as an EPA I in one trial and an EPA II in the second; formulation 1079 was categorized as an EPA III in one trial and an EPA IV in the second. A fourth material also had two sets of animal data reported, but this material is not specifically indicated since both of the animal tests predicted the same EPA category.

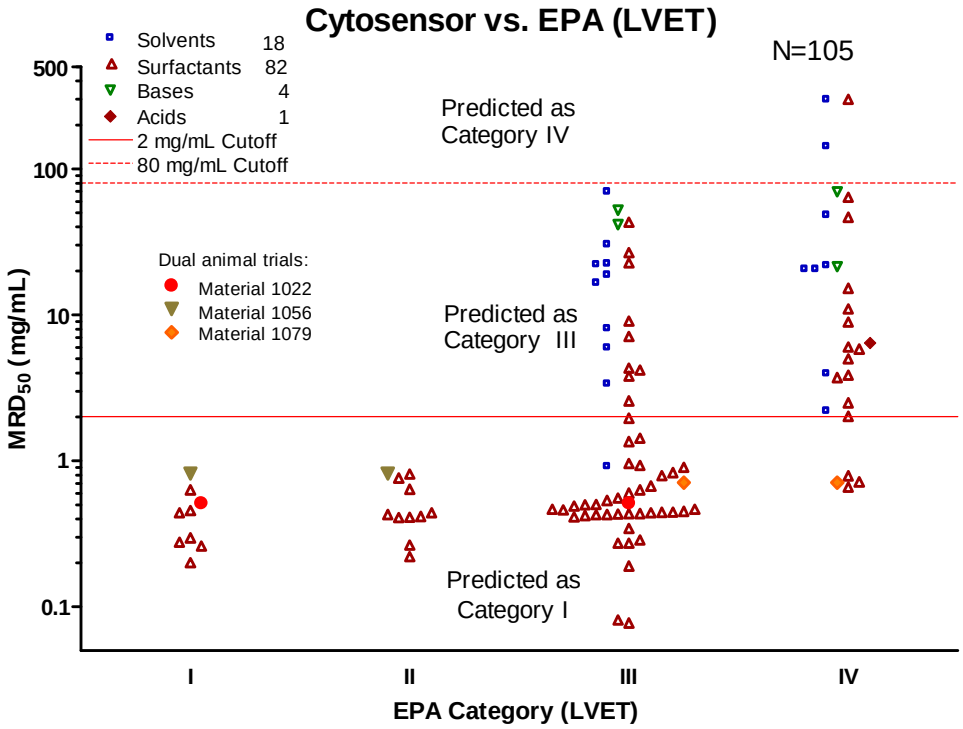


Figure 6-1 Cytosensor MRD_{50} values plotted against EPA toxicity categories determined by the LVET. Suggested cut-off values with their predicted EPA categories are included. There are 105 unique materials; however, 3 materials are graphed with 2 different EPA categories since they were tested twice in the animal trials with different results each time.

The following contingency table (Table 6-2) gives an analysis of the performance based on the cut-offs shown in Figure 6-1. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose EPA toxicity category is greater than III. One hundred percent of the animal test determined EPA Categories I and II were captured by this model. There were no underpredictions of Category I or II materials. In addition there were no underpredictions of Category III materials; all were predicted as Category III or higher. The discordant results for the CM assay and EPA toxicity categories are shown in Table 6-3. There were no underpredictions of the EPA category for any material; however, 39% of solvents and 78% of surfactants were overpredicted.

What occurs as a consequence of the conservative cut-offs is that many materials are overpredicted relative to their toxicity category as determined by the

animal test. All of the Category II materials are overpredicted as Category I's, and 67% of the Category III materials are overpredicted as Category I's. Since the CM can't distinguish between Category I and Category II, Category I is assumed as the worst case for materials with MRD_{50}'s < 2 mg/ml. Eighty-nine percent of the Category IV materials are overpredicted as Category III (75%) or I (14%) materials.

Table 6-2 Contingency table depicting the accuracy and predictivity of the CM assay for EPA toxicity categories (determined by positive responses in the LVET) using cut-off values of MRD_{50} > 80 mg/mL = IV, 80 mg/mL >MRD_{50} > 2 mg/mL = III, and MRD_{50} < 2 mg/mL = I. The model does not propose to differentiate between EPA Category I and II materials. The total number of materials is listed as 108 since the three materials with differing repeat animal scores were each scored twice.

LVET- Determined EPA Category	CM Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	9	0	0	9	100%	NA	0%
II	11	0	0	11	0%	100%	0%
III	40	20	0	60	33%	67%	0%
IV	4	21	3	28	11%	89%	NA
Total	64	41	3	108	30%		
Predictivity	14%	49%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	86%	51%	NA				

The practical advantage of such a model is that the very low irritating materials (Category III's and IV's) can be easily identified and an appropriate toxicity category applied. This will clearly result in some over labeling (75% of animal-determined IV's will be over labeled as III's), but the participating companies have accepted that this degree of over labeling will occur. The EPA appears to concur with this type of approach since the EPA label Review Manual (2003) states (for primary eye irritation of Category IV) that "…the registrant may choose to use Category III labeling."

An additional analysis was conducted to compare the performance of the prediction model with each of the different product formulation types. Table 6-3 presents the under and overpredictions associated with each product type. It can be seen that none of the product types was underpredicted. The surfactants had the highest over prediction rate (78%), however the sample size for the other product classes, especially the acids and bases, was probably too low to make a meaningful comparison.

Table 6-3 Prediction results for the CM assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Bases	Acids
Under predicted	0	0	0	0
Correctly Predicted	11 (61%)	19 (22%)	2 (50%)	0
Over Predicted	7 (39%)	66 (78%)	2 (50%)	1 (100%)

GHS Labeling Categories

A similar exercise to that shown for developing an EPA category prediction model was conducted using GHS toxicity categories. Figure 6-2 shows the CM MRD_{50}'s plotted against LVET-determined GHS categories. It can be seen that a much different pattern results with a greater number of formulations classified as non irritating in the GHS system as compared to the number that fall into the EPA non irritating category of IV's. As a result, the cut-off between NI materials and the 2B and higher categories was lowered to 10 mg/ml. The next lower cut-off to identify strongly irritating (GHS 1) materials could be set conservatively at 2 mg/ml, the same as was done for the EPA classification. Because of the overlap of MRD_{50} values for category 1 and 2A materials, no cut-off is proposed to separate these two groups. Thus materials with MRD_{50}'s <2.0 mg/ml will be categorized as 1's, those with MRD_{50}'s ≥ 2.0 and <10 mg/mL will be categorized as 2B's, and those materials with MRD50's ≥ 10 mg/mL will be categorized as 2A's.

Again materials with two sets of animal data are also indicated on the graph. Three of four replicated materials had differing GHS categories depending on the animal study used. Each of the three differed by one category between the two trials.

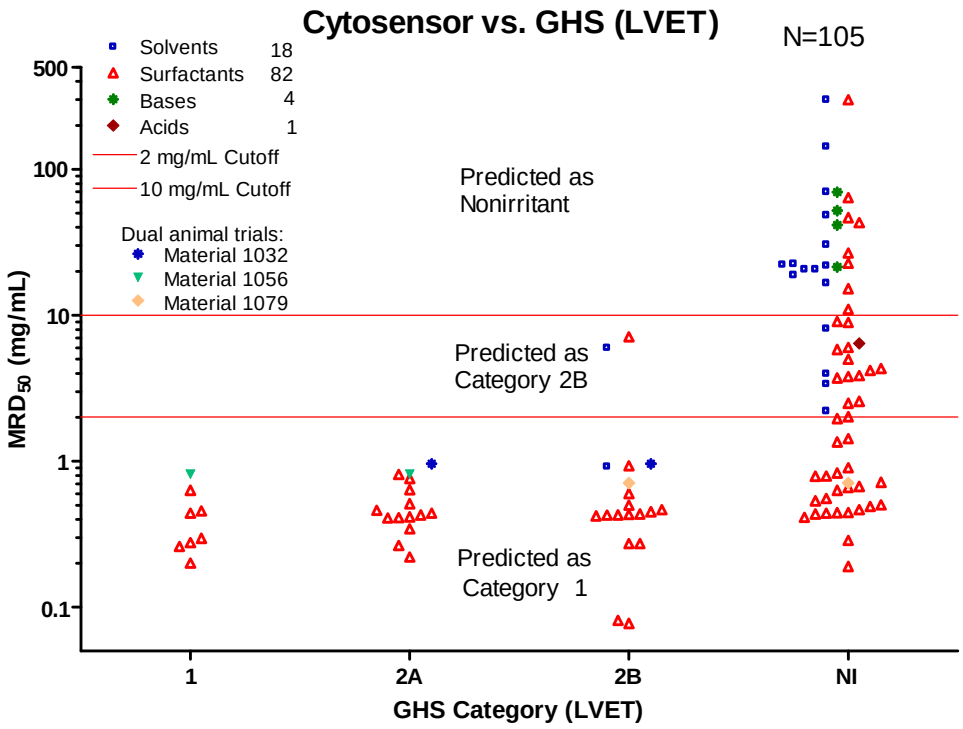


Figure 6-2 Cytosensor MRD_{50} values plotted against GHS toxicity categories determined by the LVET. All materials except oxidizing formulations are graphed. Suggested cut-off values with their predicted GHS categories are included. There are 105 unique materials; however, 3 materials have 2 GHS categories each since they were tested twice in the animal trials.

The following contingency table (Table 6-4) gives an analysis of the performance based on the cut-offs shown in Figure 6-2. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose GHS toxicity category is greater than 2B. One hundred percent of the animal test-determined GHS Categories 1 and 2A were captured by this model. There were no underpredictions of Category 1 or 2A materials. In addition there were no underpredictions of Category 2B materials; all were predicted as Category 2B or higher. The discordant results for the CM assay and GHS toxicity categories are shown in Table 6-5. There were no underpredictions of the GHS category for any material; however, 28% of solvents and 80% of surfactants were overpredicted.

What occurs as a consequence of the conservative cut-offs is that many materials are overpredicted relative to their toxicity category as determined by the animal test. All of the Category 2A materials are overpredicted as Category 1's, and 89% of the Category 2B materials are overpredicted as Category 1's. Since the CM can't distinguish between Category 1 and Category 2A, Category 1 is assumed as the worst case for materials with MRD_{50}'s < 2 mg/ml. Sixty-four percent of the Nonirritant materials are overpredicted as Category 2B (27%) or 1 (36%) materials.

Table 6-4 Contingency table depicting the accuracy and predictivity of the CM assay for GHS toxicity categories (determined by positive responses in the LVET) using cut-off values of MRD_{50} ≥ 10 mg/mL = NI, 10 mg/mL >MRD_{50} ≥ 2 mg/mL = 2B, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify GHS Category 2A materials. The total number of materials is listed as 108 since the three materials with differing repeat animal scores were each scored twice.

LVET- Determined GHS Category	CM Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2B	NI	Total			
1	8	0	0	8	100%	NA	0%
2A	15	0	0	15	0%	100%	0%
2B	17	2	0	19	11%	89%	0%
NI	24	18	24	66	36%	64%	NA
Total	64	20	24	108	31%		
Predictivity	13%	10%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	88%	90%	NA				

Over and under predictions by formulation type

An additional analysis was conducted to compare the performance of the prediction model with each of the different product formulation types. Table 6-5 presents the under and overpredictions associated with each product type. It can be seen that none of the product types was underpredicted. The surfactants had the highest over prediction rate (80%), however the sample size for the other product classes, especially the acids and bases, was probably too low to make a meaningful comparison.

Table 6-5 Number of discordant results (and percentages) for the CM assay and GHS toxicity categories.

	Solvents	Surfactants	Bases	Acids
Under predicted	0	0	0	0
Correctly Predicted	13 (72%)	17 (20%)	4 (100%)	0
Over Predicted	5 (28%)	68 (80%)	0	1 (100%)

6.1.1.1 Secondary analysis of acidic and alkaline materials

The first pass analysis described above utilized all of the submitted materials (with the exception of oxidizing formulations) for which adequate animal data were available to determine an EPA or GHS category. However, there has always been some concern that the CM should not be used for acidic or alkaline materials (pH ≤ 4.0 or ≥10.0). Therefore, we conducted a second analysis in which materials fitting

the above acid or alkaline definitions (or for which one of the 3 product activity categories was described as acid or alkaline) were omitted from the database.

Table 6-6 describes the seventeen materials identified as fitting the description as acid or alkaline. It can be seen that all of the materials were EPA Category III or IV materials and that none of the materials were underpredicted by the CM assay as might be hypothesized from the operation of the CM (cells exposed to an increasing dilution series of the test material which might quickly change the pH).

Table 6-6 Distribution of EPA categories for the 17 materials from the CM database classified as acid or alkaline.

	CM-defined EPA Category		
LVET-defined EPA Category	**I**	**III**	**IV**
III	1	7	0
IV	0	9	0

Figure 6-3 shows the distribution of MRD_{50} values for the non-acidic, non-alkaline materials plotted against EPA labeling categories (determined by the LVET). Even with the seventeen acidic/alkaline materials removed, there is not a significant change in the distribution among EPA determined categories. The same cut-off values as determined for Figure 6-1 were used.

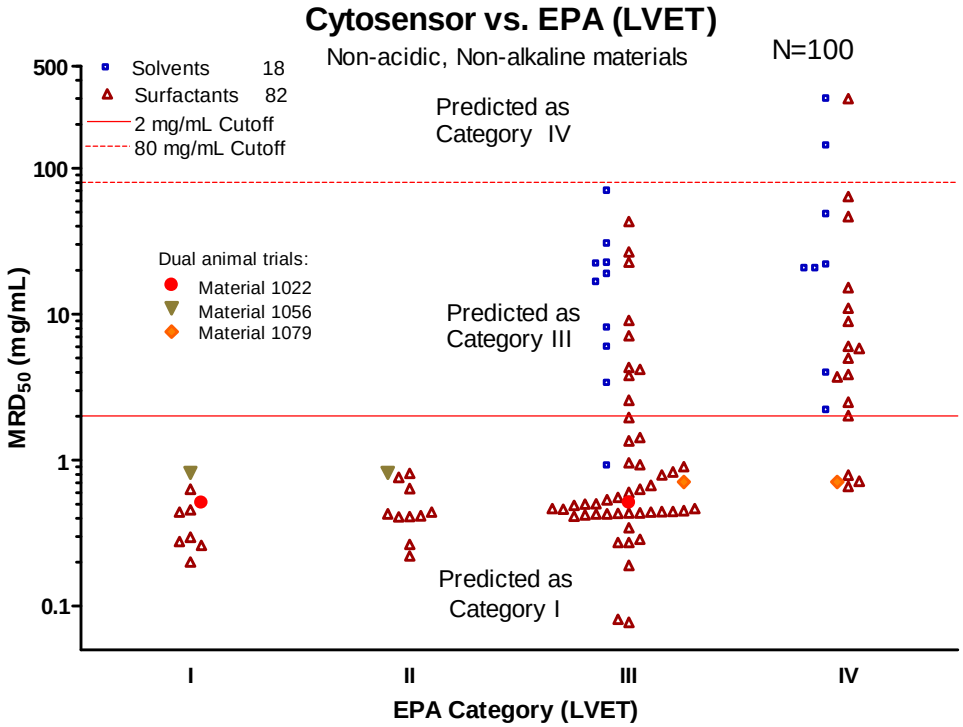


Figure 6-3 Cytosensor MRD_{50} values plotted against EPA toxicity categories determined by the LVET. Only non-acidic, non-alkaline materials are graphed. Suggested cut-off values with their predicted EPA categories are included. There are 100 unique materials; however, 3 materials have 2 values since they were tested twice in the animal trials.

As expected from the results shown in Table 6-7, the performance of the CM assay is very similar when the acid and alkaline materials are removed (Table 6-7 versus Table 6-2). There are still no underpredictions and while positive predictive values increase somewhat, the concordance decreases somewhat (due to the removal of 2 Category III materials which were correctly predicted by the CM). Thus we do not feel that acid or alkaline materials need to be excluded from analysis by the Cytosensor and propose to keep them in the applicability domain for the CM assay. The discordant results for the CM assay and EPA toxicity categories without acid/alkaline materials are shown in Table 6-8. There were no underpredictions of the EPA category for any material; however, 39% of solvents and 78% of surfactants were still overpredicted.

Table 6-7 Contingency table depicting the accuracy and predictivity of the CM assay for EPA toxicity categories (determined by positive responses in the LVET) of non-acidic, non-alkaline materials using cut-off values of MRD_{50} ≥ 80 mg/mL = IV, 80 mg/mL >MRD_{50} ≥ 2 mg/mL = III, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify EPA Category II materials.

LVET- Determined EPA Category	CM Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	9	0	0	9	100%	NA	0%
II	11	0	0	11	0%	100%	0%
III	40	18	0	58	31%	69%	0%
IV	4	18	3	25	12%	88%	NA
Total	64	36	3	103	29.1%		
Predictivity	14.1%	50%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	85.9%	50%	NA				

A similar exercise was conducted using GHS toxicity categories. Figure 6-4 shows the CM MRD_{50}'s plotted against LVET-determined GHS categories with the seventeen acidic/alkaline materials removed. Even with the seventeen acidic/alkaline materials removed, there is not a significant change in the distribution among GHS determined categories. The same cut-off values as determined for Figure 6-2 were used.

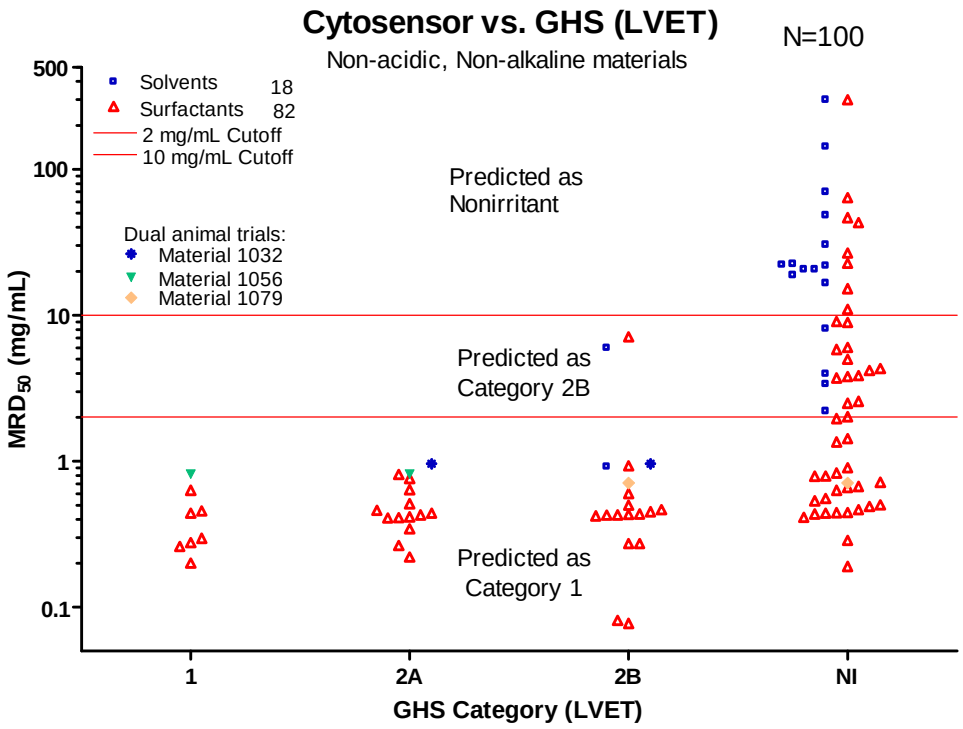


Figure 6-4 Cytosensor MRD_{50} values plotted against GHS toxicity categories determined by the LVET. Only non-acidic, non-alkaline materials are graphed. Suggested cut-off values with their predicted GHS categories are included. There are 100 unique materials; however, 3 materials have 2 values since they were tested twice in the animal trials.

As expected from the results shown in Table 6-9, the performance of the CM assay is very similar when the acid and alkaline materials are removed (Table 6-9 versus Table 6-4). There are still no underpredictions and while positive predictive value increase somewhat, the concordance decreases somewhat (due to the removal of 4 Nonirritant materials which were correctly predicted by the CM). Thus we do not feel that acid or alkaline materials need to be excluded from analysis by the Cytosensor and propose to keep them in the applicability domain for the CM assay.

Table 6-8 Contingency table depicting the accuracy and predictivity of the CM assay for GHS toxicity cate8gories (determined by positive responses in the LVET) using cut-off values of MRD_{50} ≥ 10 mg/mL = NI, 10 mg/mL >MRD_{50} ≥ 2 mg/mL = 2B, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify GHS Category 2A materials.

LVET- Determined GHS Category	CM Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2B	NI	Total			
1	8	0	0	8	100%	NA	0%
2A	15	0	0	15	0%	100%	0%
2B	17	2	0	19	11%	89%	0%
NI	24	17	20	61	33%	67%	NA
Total	64	19	20	103	29%		
Predictivity	13%	11%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	87%	89%	NA				

6.1.2 Using the Draize assay to define a prediction model for the CM

Since the above analyses were conducted with EPA or GHS categories determined by the LVET, we next evaluated whether similar prediction models would have been developed if the traditional Draize test were used to obtain EPA classifications. It is known that the LVET gives somewhat lower MAS scores than does the Draize test, but the LVET is still more sensitive – and thus overpredictive – of the human response (see discussion in Section 4.7). We found two studies which used materials (surfactants and surfactant-containing formulations) similar to those which are the focus of this BRD (anti-microbial cleaning products). One of the two studies - the CTFA Phase III study - is important because it uses both LVET and Draize evaluation of surfactant-containing products; hence the results using the two methods can be directly compared for an identical set of formulations (see Table 5-4). The second study – the COLIPA study - used only the Draize test for characterization, but it contained some formulations similar to those which are being used in this BRD.

6.1.2.1 CTFA Phase III Evaluation

Previous analysis (Cytosensor BRD prepared for ECVAM) of the CTFA Phase III study indicated that the lower cut-off value to identify EPA Category I materials should be set at 2 mg/mL, identical to that which we have proposed in the preceding analysis of the data submitted specifically for this BRD. There are some differences in the chemical makeup of the two data sets, but they do overlap considerably in the type of chemical formulation, both data sets being highly biased

towards surfactant-based formulations. The distribution of product categories submitted with CTFA Phase II *in vitro* and Cytosensor data is shown in Table 6-11.

Table 6-9 Distribution of product categories originally submitted with both animal eye irritation data and CTFA Phase III *in vitro* data.

Product Categories	Number of products tested
Oxidizers	0
Surfactants	25
Acids	0
Bases	0
Solvents	0
Total	25

Further analysis of this study brings up the importance of being aware of the variability of the animal test in making EPA toxicity category decisions. Figure 6-5 shows that there are 2 Category I materials (identified as Facial Cleaning Foam and Gel Cleanser) which would be identified as being underpredicted (relative to the Draize classifications) by the CM assay (they both have MRD_{50} values >2 mg/mL). In order to begin to understand these apparent underpredictions, the individual animal scores for both six-rabbit tests were examined. Since the EPA currently accepts the results from three rabbit tests, we parsed the 6- rabbit test data into 20 unique, but equally likely, subgroups of three rabbit results. EPA grading criteria were then applied to each of the three-rabbit subgroups and an EPA toxicity Category determined (see Table 4-16). For the Gel Cleanser, even though the six-rabbit calculation gave a Category I result, only ten (of twenty) three-rabbit subgroups received a score of Category I; the other ten received a score of Category III. The same results were found for the Facial Cleaning Foam; ten three-rabbit subgroups received a score of Category I, and the other ten received a score of Category III. Thus if the test were performed repeatedly on the two materials using today's three-rabbit test standard, 50% of the time the materials would be graded as Category III and 50% of the time they would be graded as Category I - a difference of 2 toxicity classification grades! Thus it is extremely hard to say that the CM truly underpredicts the irritation potential of these two materials.

An additional insight from the CTFA Phase III study is the apparent over classification of the surfactant-based personal care products relative to their intended use (often on the face and around the eyes). A large number of these commonly used personal care products fall into EPA Category I (10 out of 25) when they are tested using the Draize test; however, they are categorized somewhat lower, and possibly more realistically when using the LVET.

It appears from an examination of the Figure 6-5 and Tables 6-12 & 6-14 that a decision on where to place the cut-off values would be very similar whether the LVET or the Draize data were used as the basis.

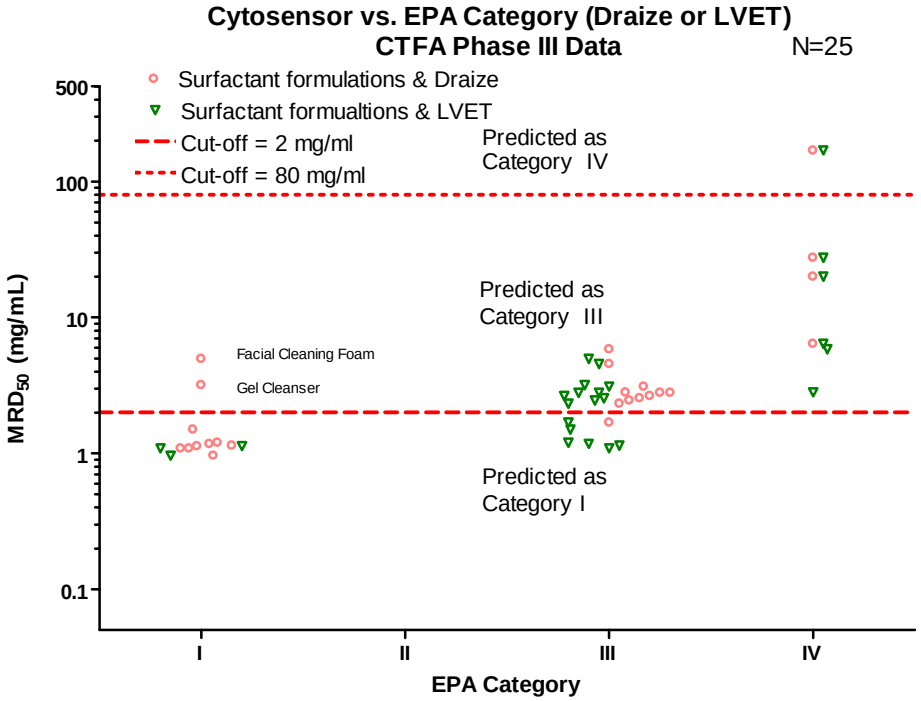


Figure 6-5 Plot of CM data versus both LVET- and Draize-defined EPA Categories for the 25 surfactant-based personal care products tested in the CTFA Phase III (Gettings, Lordo et al. 1996) evaluation using cut-off values of MRD_{50} ≥ 80 mg/mL = IV, 80 mg/mL >MRD_{50} ≥ 2 mg/mL = III, and MRD_{50} < 2 mg/mL = I. The model does not propose to identify EPA Category II materials.

The following contingency tables (Table 6-12 & 6-14) give an analysis of the performance based on the cut-offs shown in Figure 6-5 for the LVET-determined EPA category or the Draize-determined EPA category, respectively. One hundred percent of the LVET-determined EPA Category I materials were captured by this model; however, 20% of the Draize-determined EPA Category I materials were underpredicted by the CM. In contrast, 38% of LVET-determined EPA Category III materials were overpredicted, whereas, only 9% of Draize-determined EPA Category III materials were overpredicted. The discordant results for the CM assay and EPA toxicity categories are shown in Table 6-13 & 6-15. There were no underpredictions of the LVET-determined EPA category, but 8% of Draize-determined EPA category was underpredicted. There was a significant amount of overprediction for both LVET and Draize–determined EPA categories mainly due to the EPA Category IV materials being overpredicted as Category III.

Table 6-10 Contingency table presenting the accuracy and predictivity of the CM for EPA toxicity categories (LVET-determined) for the 25 surfactant-based personal care products in the CTFA Phase III study (Gettings, Lordo et al. 1996).

LVET- Determined EPA Category	LVET Category Predicted by CM				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	3	0	0	3	100%	NA	0%
II	0	0	0	0	0%	0%	0%
III	6	10	0	16	63%	38%	0%
IV	0	5	1	6	17%	83%	NA
Total	9	15	1	25	56%		
Predictivity	33%	67%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	67%	33%	NA				

Table 6-11 shows the overall results based on product category – in this instance for surfactants only.

Table 6-11 Discordant results for the CTFA CM study and EPA toxicity categories (LVET-determined).

	Surfactants
Under predicted	0
Correctly Predicted	14
Over Predicted	11

Table 6-12 Contingency table presenting the accuracy and predictivity of the CM for EPA toxicity categories(Draize-determined) for the 25 surfactant-based personal care products in the CTFA Phase III study (Gettings, Lordo et al. 1996).

Draize-Determined EPA Category	Draize Category Predicted by CM				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	8	2	0	10	80%	NA	20%
II	0	0	0	0	0%	0%	0%
III	1	10	0	11	91%	9%	0%
IV	0	3	1	4	25%	75%	NA
Total	9	15	1	25	76%		
Predictivity	89%	67%	100%				
Category under predicted	NA	13%	0%				
Category over predicted	11%	20%	NA				

Table 6-13 shows the overall results based on product category – in this instance for surfactants only.

Table 6-13 Discordant results for the CTFA CM study and EPA toxicity categories (Draize-determined).

	Surfactants
Under predicted	2 (8%)
Correctly Predicted	19 (76%)
Over Predicted	4 (16%)

6.1.2.2 COLIPA Evaluation

The distribution of product categories for the COLIPA *in vitro* and Cytosensor data is shown in Table 6-14. The COLIPA evaluation was for surfactant and surfactant-containing materials only.

Table 6-14 Distribution of product categories originally submitted with both animal eye irritation data and COLIPA *in vitro* data.

Product Categories	Number of products tested
Oxidizers	0
Surfactants	19
Acids	0
Bases	0
Solvents	0
Total	19

Figures 6-6 & 6-7 show MRD_{50} scores obtained in the COLIPA evaluation of *in vitro* assays for eye irritation. The cut-off values for MRD_{50} scores have been empirically chosen to identify, where possible, the various toxicity categories. In attempting to select cut-off values we first tried those that were chosen from the CTFA Phase III studies (see preceding sections). Since these appeared adequate, we continued the analysis with these values for the sake of consistency. As with the CTFA Phase III studies, in the case of the GHS system and the EPA system which have 4 categories, the overlap of MRD_{50} response was so large that it was deemed impossible to differentiate between the two middle categories (either EPA II and III or GHS 2A and 2B) from each other. This analysis was made even more difficult because of the distribution of the toxicity classifications. There were only two GHS Draize determined 2A or 2B materials. Hence only upper (to possible identify non-irritants) and lower (to possibly identify severe irritants) cut-off values are shown.

For the COLIPA GHS data set (Figure 6-6), it appeared a cut-off value of >10 mg/mL might be appropriate to identify the GHS nonirritants from the more irritating materials while a higher cut-off of 80 mg/ml seemed appropriate to use with the EPA classifications. The cut-off of <2 mg/ml was retained for identifying both GHS 1 or EPA I materials. However, as seen in most of the previous analyses, there

were very few materials in the GHS 2A or 2B, or EPA II categories. This makes it difficult to determine exactly where the cut-off between these intermediate irritating categories and the mild categories lies. Additionally the EPA classification had only two Category IV materials, again making a decision for a cut-off problematic. Hence only upper (to possible identify non-irritants) and lower (to possibly identify severe irritants) cut-off values are shown on the scatter plots. Products falling between these limits are considered Category III.

When CM MRD_{50}'s were plotted against the EPA categorization scheme (Figure 6-7), there were two Category I materials (labeled #21 and #36 on the scatter plot) that appeared to be underpredicted as Category III's. However, material #21 (1% benzalkonium chloride) has two sets of animal test data reported in the ECETOC eye irritation report (ECETOC 1992) from which the COLIPA study took its *in vivo* data. We chose to graph the highest category data, but the EPA category of the replicate animal test was a Category 3 – the same as was estimated by the Cytosensor MRD_{50}. When the second underpredicted EPA Category 1 material was decoded it was found to be a baby shampoo formulation. Thus the two EPA Category I "underpredictions" may not be as much of a concern as first suspected.

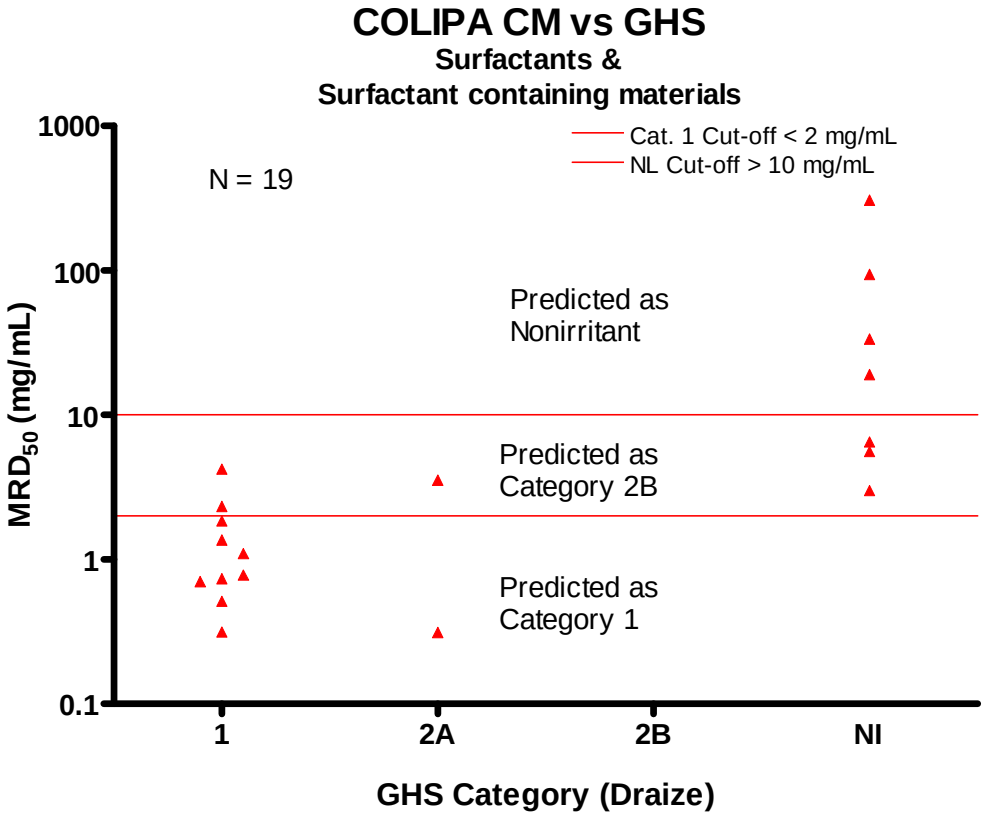


Figure 6-6 Surfactant and surfactant-containing formulation results of the COLIPA study related to GHS classification. Data points indicate the mean MRD_{50} for both laboratories (with the exception of two data points where only one laboratory made the determination). In some cases data points have been slightly offset along the X-axis in order to clearly separate them from data of similar magnitude.

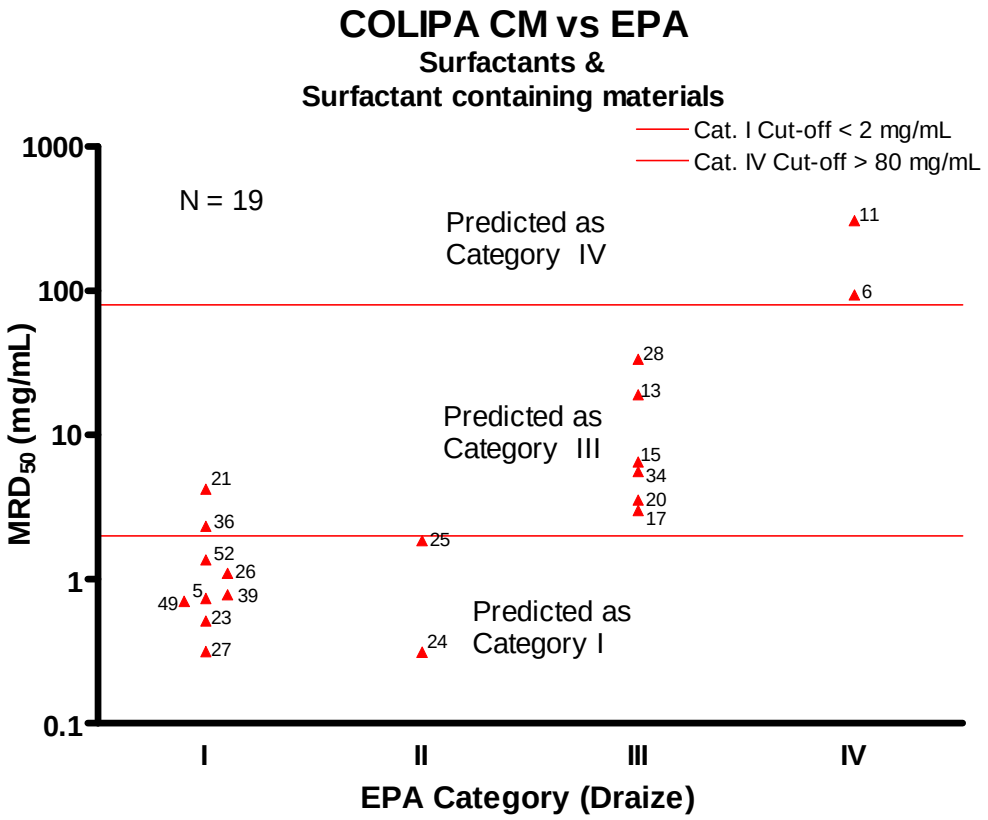


Figure 6-7 Surfactant and surfactant-containing formulation results of the COLIPA study related to EPA classification. Data points indicate the mean MRD_{50} for both laboratories with the exception of 24 and 52 which were done in one laboratory only. In some cases data points have been slightly offset along the X-axis in order to clearly separate them from data of similar magnitude. The individual materials can be identified by comparing the numbers adjacent to the symbols with the numbering code given in Table 5.3.1.3.

Contingency Tables 6-15 & 6-17 give an analysis of the performance based on the cut-offs show in Figures 6-6 & 6-7, respectively. It appears from the graphs that the CM does not have the ability to clearly separate the surfactants or surfactant-containing materials used in the COLIPA study into the four Draize test defined GHS or EPA Categories. However, severe irritants seem to be reasonably predicted when MRD_{50} scores of less than 2 are used. Using this lower cut-off value, there is a high positive predictive value for GHS Category 1 (80%; 8 of 10 materials) and EPA Category I (78%; 7 of 9 materials). There also seems to be good predictivity for EPA Category III materials and possibly for the Category IV materials as well.

Even though the positive predictive value was high using a lower cut-off of MRD_{50} <2 mg/ml, the sensitivity was lower, with several chemicals being underpredicted by at least one toxicity category by the GHS, and EPA classification system. Overpredictions of mild materials (GHS Nonirritant, and EPA IV), did not occur as often. <u>One very important conclusion from both the CTFA Phase III study and the COLIPA study is that the prediction model (cut-off values) determined for</u>

the CM using the traditional Draize assay is identical to the prediction model determined using the LVET assay.

The discordant results for the CM assay and the GHS & EPA toxicity categories are shown in Tables 6-16 & 6-18, respectively. The majority of the materials were correctly predicted with 63% correctly predicted with the GHS category and 79% correctly predicted with the EPA category. The amount of underprediction was 16% for the GHS category and 11% for the EPA category.

Table 6-15 COLIPA surfactant and surfactant containing materials. Contingency table depicting the concordance and predictivity of the CM assay for GHS toxicity classifications when the cut-off values shown in Figure 6-6 are applied.

Draize Determined GHS Category	GHS Category Predicted by CM 1	2B	NI	Total	Concordance	Toxicity Overpredicted	Toxicity Underpredicted
1	8	2	0	10	80%	NA	20%
2A	1	1	0	2	0%	50%	50%
2B	0	0	0	0	0%	0%	0%
NI	0	3	4	7	57.1%	42.9%	NA
Total	9	6	4	19	63.1%		
Predictivity	88.9%	0%	100.0%				
Category Underpredicted	NA	50%	0%				
Category Overpredicted	11.1%	50%	NA				

Table 6-16 Discordant results for the COLIPA CM study and GHS toxicity categories.

	Surfactants
Under predicted	3 (16%)
Correctly Predicted	12 (63%)
Over Predicted	4 (21%)

Table 6-17 COLIPA surfactant and surfactant containing materials - Contingency table depicting the concordance and predictivity of the CM assay for EPA toxicity classifications when the cut-off values shown in Figure 6-7 are applied.

Draize Determined EPA Category	EPA Category Predicted By CM I	III	IV	Total	Concordance	Toxicity Overpredicted	Toxicity Underpredicted
I	7	2	0	9	77.8%	NA	22.2%
II	2	0	0	2	0%	100%	0%
III	0	6	0	6	100%	0%	0%
IV	0	0	2	2	100%	0%	NA
Total	9	8	2	19	78.9%		
Predictivity	77.8%	75%	100%				
Category Underpredicted	NA	25%	0%				
Category Overpredicted	22.2%	0%	NA				

Table 6-18 Discordant results for the COLIPA CM study and EPA toxicity categories.

	Surfactants
Under predicted	2 (11%)
Correctly Predicted	15 (78%)
Over Predicted	2 (11%)

6.1.3 Cytosensor studies without animal data

Many companies do not currently conduct rabbit eye irritation tests on cleaning products; hence, many *in vitro* study data were submitted to this BRD without accompanying animal data. We compared the distribution of these scores using the same cut-off values for classification that were used in the analysis of the predictive capacity of the CM for LVET-determined EPA toxicity Categories (see, for example, Table 6-2). Figure 6-8 shows the distribution of CM scores for the products without animal data. Using the previous suggested cut-offs ($MRD_{50} \geq 80$ mg/mL = IV, 80 mg/mL $> MRD_{50} \geq 2$ mg/mL = III, and $MRD_{50} < 2$ mg/mL = I), 1.9% of the materials would be Category IV's, 24.7% would be Category III's and 73.5% would be Category I. This compares to the products with paired animal and CM data analyzed in Table 6-2 where the materials assigned to categories by CM scores were 2.8% Category IV's, 38% Category III's and 59% Category I's. Thus it appears that the distribution of CM-measured toxicities for the set of materials used to determine cut-off values (those which were tested with both the rabbit test and the CM test) were somewhat less irritating than those which were tested in the CM alone.

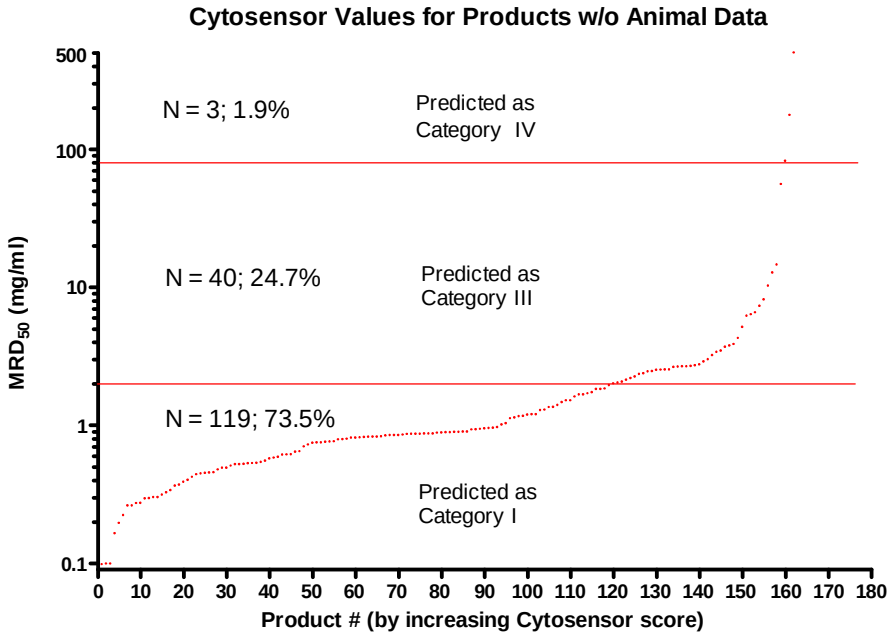


Figure 6-8 Distribution of CM scores for the products without animal data using cut-offs of $MRD_{50} \geq 80$ mg/mL = IV, 80 mg/mL $> MRD_{50} \geq 2$ mg/mL = III, and $MRD_{50} < 2$ mg/mL = I.

6.1.4 Conclusion for the Cytosensor assay

The Cytosensor assay appears to be most useful at the less irritating portion of the irritation spectrum. It is capable of identifying both Category III and IV materials, although most Category IV materials will be overpredicted as Category III materials. None of the 105 materials cleaning products were under predicted for LVET-defined EPA toxicity categories. Over predictions were much more frequent, but this was driven by the fact that the CM assay seems incapable of clearly differentiating between Category I and Category II materials. In fact many Category III materials (67%) were also over predicted as EPA Category I. The corporate participants have agreed that this outcome of over labeling some materials is acceptable to them.

Similar results were found with the prediction model for GHS categories, with the exception that 36% of the GHS Non-irritating materials were clearly identified as such by the CM. However only 11% of the GHS category 2B materials were correctly identified; the rest were over predicted as Category 1.

The corporate participants in this program have agreed that the outcome is acceptable to them. Another assay (we propose BCOP) will be used as a second tier test to differentiate EPA Category I from EPA Category II (and lower) materials, if needed.

Again it is important to note that the prediction model for both the GHS and EPA toxicity categories is the same whether determined by the Draize assay or the LVET assay.

Historical knowledge of the performance of the Cytosensor assay plus the preceding analysis of the Cytosensor data in this BRD have led us to the following recommendations:

1) **Anti-microbial cleaning products having an oxidizing chemistry should not be tested with the Cytosensor assay.**

2) **Only fully water soluble anti-microbial cleaning products can be tested with the Cytosensor assay.**

3) **If the anti-microbial cleaning product has an MRD_{50} score of <2 mg/ml, it is classified as EPA Category I or GHS Category 1.**

4) **If the anti-microbial cleaning product has an MRD_{50} score of ≥2 mg/ml, but < 80 mg/ml, it is classified as EPA Category III. If the anti-microbial cleaning product has an MRD_{50} score of ≥2 mg/ml, but <10 mg/ml, it is classified as GHS Category 2B.**

5) **If the anti-microbial cleaning product has an MRD_{50} score of ≥80 mg/ml, it is classified as EPA Category IV. If the anti-microbial**

cleaning product has an MRD_{50} score of ≥10 mg/ml, it is classified GHS Category NI.

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA I or GHS 1 is actually an EPA II or a GHS 2A, it should be further tested in the BCOP assay.**

6.2 EpiOcular predictive capacity

6.2.1 Company submissions

EPA Labeling Categories (LVET-determined)

Table 6-21 gives the distribution of product categories originally submitted with both animal eye irritation data (LVET) and EpiOcular *in vitro* data. This distribution is more highly weighted to formulations having oxidizing chemistry than is the total data submitted by participating companies for all of the other *in vitro* tests.

Table 6-19 Distribution of product categories originally submitted with both animal eye irritation data (LVET) and EpiOcular data.

Product Categories	Number of products tested
Oxidizers	9
Surfactants	12
Solvents	4
Total	25

Figure 6-9 shows the full distribution of ET_{50} values for all of the 25 materials for which data were available when plotted against EPA labeling categories (determined by the LVET). EPA categories are not equally represented since only one Category II material and three Category I materials are present. This is not surprising since this method was not intended for identifying more severe irritants. It is immediately apparent from Figure 6-9 that the distribution of ET_{50} scores across the EPA labeling categories is not random. EPA Category I materials have ET_{50}'s <4 min, while most EPA Category III and IV materials have ET_{50}'s > 10 min. This distribution allowed us to set cut-off values by eye for predicting EPA labeling categories. We attempted to choose cut-offs conservatively with a bias towards having as few under predictions as was reasonable. No statistical methods were employed to construct the proposed prediction model. Thus – for this somewhat limited data set - all materials in the highest toxicity category can be identified with a cut-off value of 4 min. However, a number of the Category III and IV materials also fall below this ET_{50} value.

ET_{50} values for EPA Category III and IV materials have significant overlap. Thus it is not possible from this data set to suggest an upper cut-off limit to separate EPA Category IV materials from Category III materials. However, materials having ET_{50} values above ~ 70 min would likely be Category IV materials. The consequence of this is that many EPA Category III and a few EPA Category IV materials would be overpredicted. Figure 6-9 shows a plot of ET_{50} values versus EPA category classification (by LVET) with the above proposed cut-off values added.

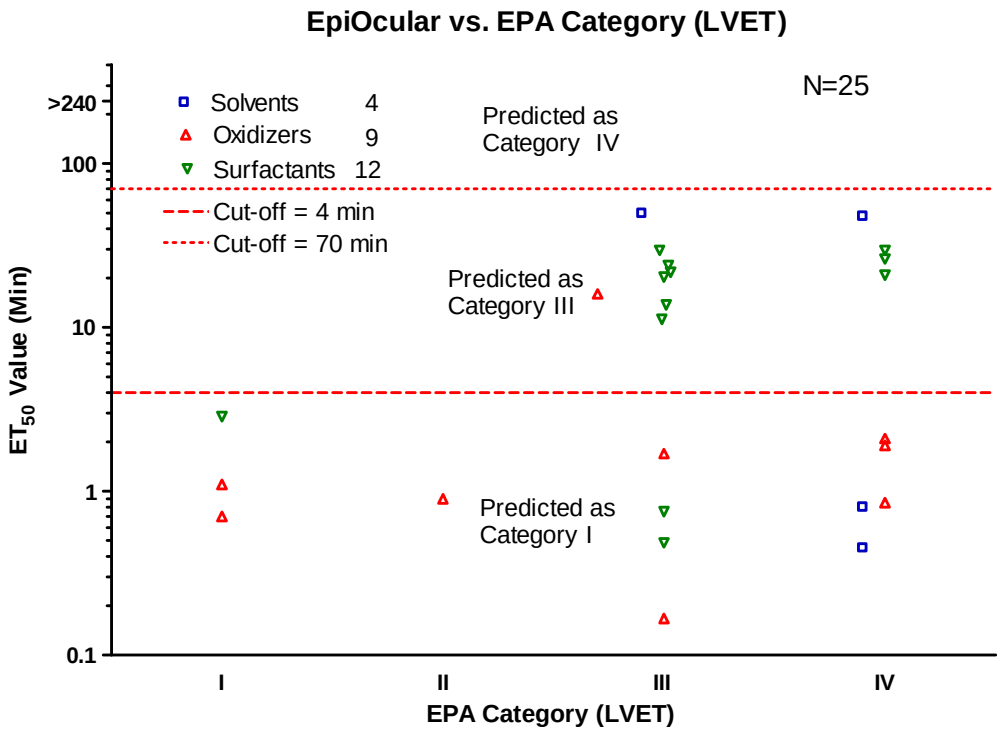


Figure 6-9 EpiOcular ET_{50} values plotted against EPA categories determined by the LVET. Suggested cut-off values with their predicted EPA categories are included.

The contingency table (Table 6-20) gives an analysis of the performance based on the cut-offs shown in Figure 6-9. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose EPA toxicity category is greater than III. There was a significant amount of overprediction for EPA Category IV materials (100%); however, there were no underpredictions for any of the EPA categories.

Table 6-20 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = IV, and ET_{50} < 4 min = I. ET_{50} values ≥4 min and <70 min are predicted to be EPA III. The model does not propose to identify EPA Category II materials.

LVET- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	3	0	0	3	100%	NA	0%
II	1	0	0	1	0%	100%	0%
III	4	8	0	12	67%	33%	0%
IV	5	4	0	9	0%	100%	NA
Total	13	12	0	25	44%		
Predictivity	23%	67%	0%				
Category under predicted	NA	0%	0%				
Category over predicted	77%	33%	NA				

An additional analysis was conducted to compare the performance of the prediction model with each of the different product formulation types. Table 6-21 presents the under and overpredictions associated with each product type. It can be seen that none of the product types was underpredicted; however, 75% of solvents, 42% of surfactants, and 67% of oxidizers were overpredicted

Table 6-21 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses)..

	Solvents	Surfactants	Oxidizers
Under predicted	0	0	0
Correctly Predicted	1 (25%)	7 (58%)	3 (33%)
Over Predicted	3 (75%)	5 (42%)	6 (67%)

It appears that almost all of the oxidizing formulations (8 out of 9) are predicted to be Category I materials by the EpiOcular assay, even though their *in vivo* irritation potential appears to vary considerably (from Category IV to Category I) in the animal test. This may be a reflection of the epithelial-only nature of the EpiOcular tissue. In this model, it may be possible for the oxidizing formulations to kill almost all of the EpiOcular tissue *in vitro* (and thus have the highest score possible which would be a Category I), while *in vivo* the material might penetrate only a small way past the epithelium into the stroma and thus cause a toxicity that would be a Category III or at the most a Category II. *Because of these significant (all of the over predictions were by at least 2 toxicity categories) and consistent overpredictions, we suggest that oxidizers be tested only in the BCOP assay.* Figure 6-10 shows the distribution of full ET_{50} values for all of the 16 materials for which data were available when plotted against EPA labeling categories (determined by the LVET) without the oxidizers.

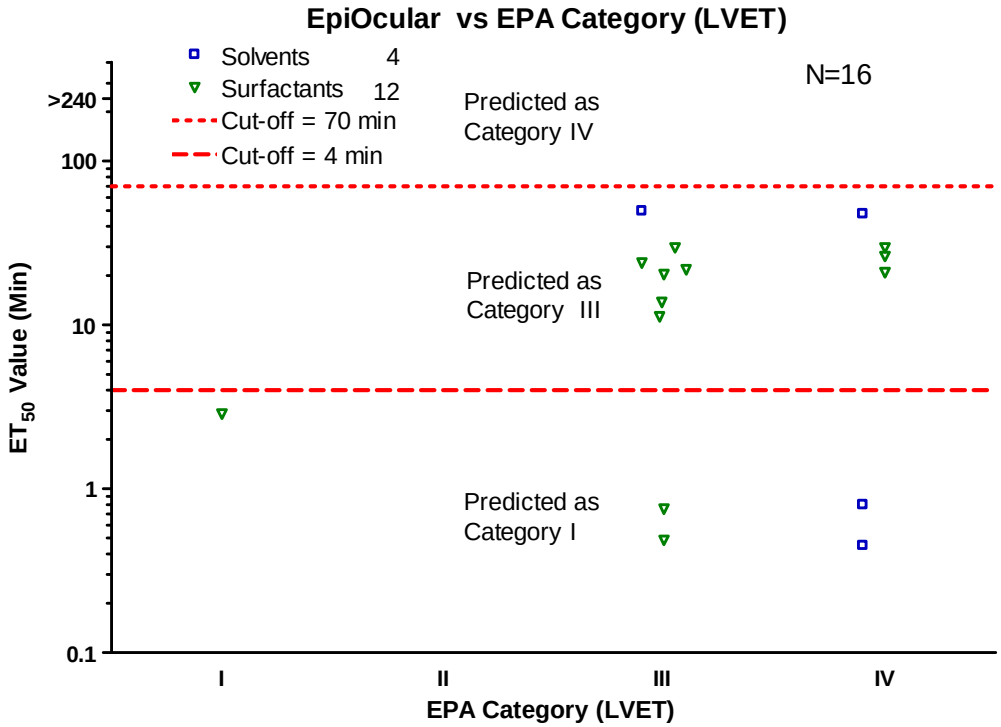


Figure 6-10 EpiOcular ET_{50} values plotted against EPA categories determined by the LVET. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted EPA categories are included.

The contingency table (Table 6-24) gives an analysis of the performance based on the cut-offs shown in Figure 6-10. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose EPA toxicity category is greater than III with the caveat that no Category II materials were available for this analysis. All animal test-determined EPA Category I formulations were captured by this model. There were no underpredictions of Category I materials. In addition there were no underpredictions of Category III materials; all were predicted as Category III or higher.

What occurs as a consequence of the conservative cut-offs is that many materials are overpredicted relative to their toxicity category as determined by the animal test (LVET). Twenty-two percent of the Category III materials are overpredicted as Category I's, and 100% of the Category IV materials are overpredicted as Category III or I materials.

Table 6-22 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = ET_{50} values ≥4 min and <70 min are predicted to be EPA III IV, and ET_{50} < 4 min = I. ET_{50} values ≥4 min and <70 min are predicted to be EPA III. The model does not propose to identify EPA Category II materials.

LVET- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	1	0	0	1	100%	NA	0%
II	0	0	0	0	0%	0%	0%
III	2	7	0	9	78%	22%	0%
IV	2	4	0	6	0%	100%	NA
Total	5	11	0	16	50%		
Predictivity	20%	64%	0%				
Category under predicted	NA	0%	0%				
Category over predicted	80%	36%	NA				

The discordant results for the EpiOcular assay and EPA toxicity categories without oxidizers are shown in Table 6-23. Because the prediction model was not changed, the results for the solvents and surfactants remain the same as in Table 6-21.

Table 6-23 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants
Under predicted	0	0
Correctly Predicted	1 (25%)	7 (58%)
Over Predicted	3 (75%)	5 (42%)

The practical advantage of such a model is that the very low irritating materials (Category III's and IV's) can be identified and an appropriate toxicity category applied. This will clearly result in some over labeling (67% of animal-determined IV's will be over labeled as III's and 33% as I's), but the participating companies have accepted this degree of over labeling will occur. Alternatively, all of the EO predicted Category I materials could be retested in the BCOP assay. The EPA appears to concur with this type of approach, at least for over labeling by one category, since the EPA label Review Manual (2003) states (for primary eye irritation of Category IV) that "…the registrant may choose to use Category III labeling."

GHS Labeling Categories (LVET-determined)

It can be seen from Figure 6-11 that the distribution of ET_{50} scores across the GHS labeling categories is not random. All GHS Category 1 materials have ET_{50}'s <4 min, while most GHS Category 2B and Nonirritant materials have ET_{50}'s > 10 min. Thus – for this somewhat limited data set - all materials in the highest toxicity category can be identified with a cut-off value of 4 min. However, a number of the Category 2B and Nonirritant materials also fall below this ET_{50} value.

ET_{50} values for GHS Category 2B and Nonirritant materials have considerable overlap. However, due to the limited number of Category 2B data points, it is not possible from this data set to suggest an upper cut-off limit to separate GHS Category 2B materials from Nonirritant materials. Materials having ET_{50} values above ~ 70 min would likely be Nonirritant materials. The consequence of this is that many GHS Category 2B and a few Nonirritant materials would be overpredicted.

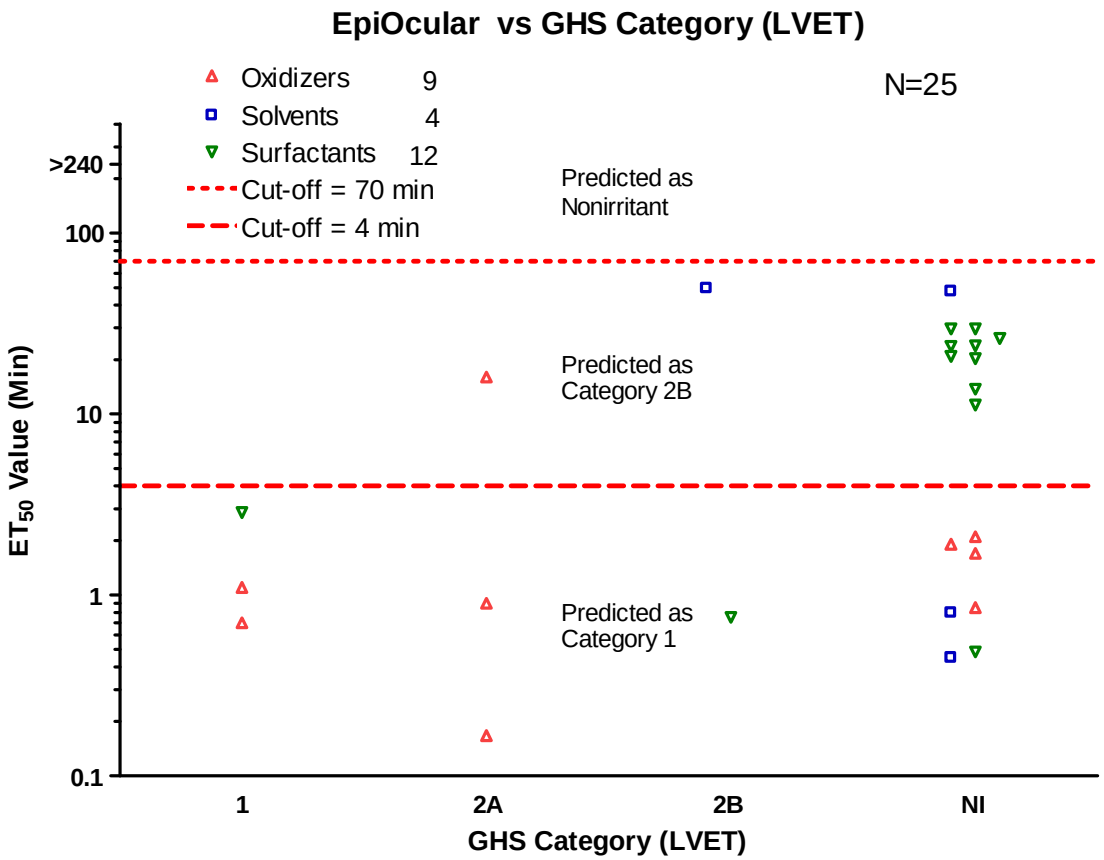


Figure 6-11 EpiOcular ET_{50} values plotted against GHS categories determined by the LVET. Suggested cut-off values with their predicted GHS categories are included.

The contingency table (Table 6-24) gives an analysis of the performance based on the cut-offs shown in Figure 6-11. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose GHS toxicity category is greater than 2B. There was a significant amount of overprediction for GHS Nonirritant materials (100%).

Table 6-24 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials.

LVET- Determined GHS Category	EpiOcular Predicted GHS Category 1	2B	NI	Total	Concordance	Toxicity over predicted	Toxicity under predicted
1	3	0	0	3	100%	NA	0%
2A	2	1	0	3	0%	67%	33%
2B	1	1	0	2	50%	50%	0%
NI	7	10	0	17	0%	100%	NA
Total	13	12	0	25	16%		
Predictivity	23%	8%	0%				
Category under predicted	NA	8%	0%				
Category over predicted	77%	83%	NA				

An additional analysis was conducted to compare the performance of the prediction model with each of the different product formulation types. Table 6-25 presents the under and overpredictions associated with each product type. It can be seen that none of the solvents or surfactants were underpredicted, but one of the oxidizers was underpredicted. However, 75% of solvents, 92% of surfactants, and 67% of oxidizers were overpredicted.

Table 6-25 Prediction results for the EO assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers
Under predicted	0	0	1 (11%)
Correctly Predicted	1 (25%)	1 (8%)	2 (22%)
Over Predicted	3 (75%)	11 (92%)	6 (67%)

It appears that almost all of the oxidizing formulations (8 out of 9) are predicted to be GHS Category 1 materials by the EpiOcular assay, even though their *in vivo* irritation potential appears to vary considerably (from Nonirritant to Category 1) in the animal test. Because of these significant and consistent overpredictions, the data set we again analyzed without the oxidizing formulations (Figure 6-12).

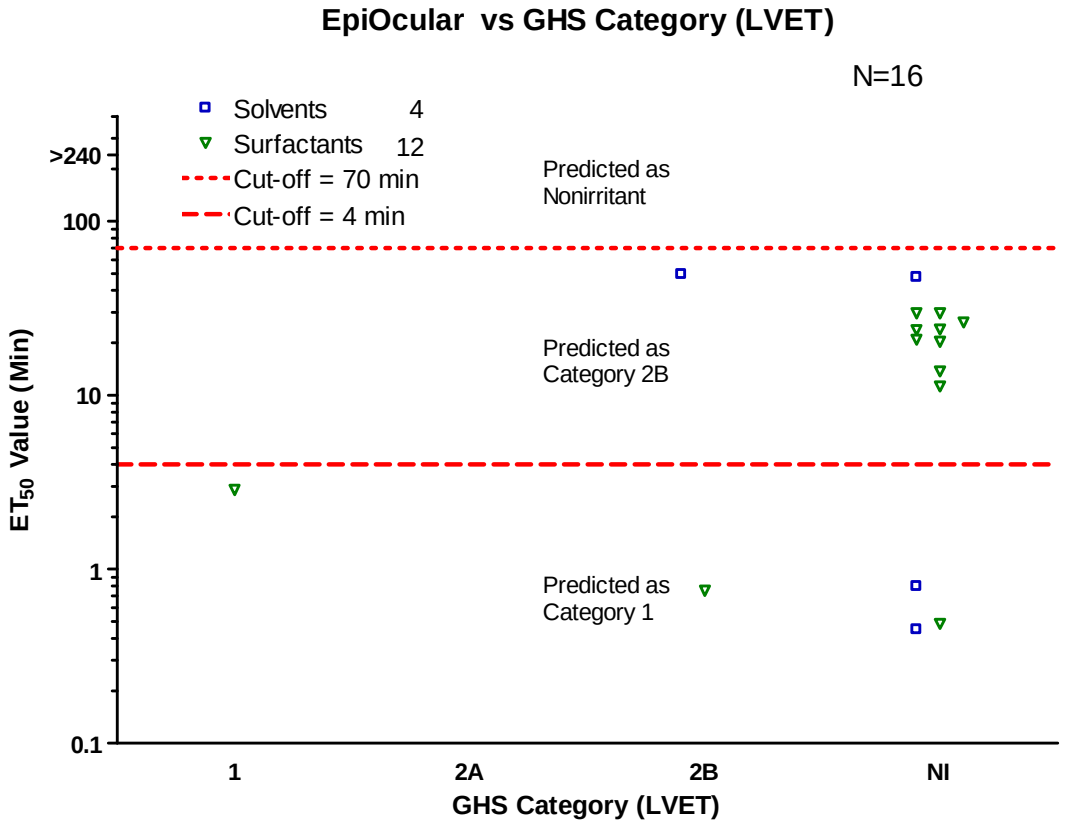


Figure 6-12 EpiOcular ET_{50} values plotted against GHS categories determined by the LVET. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted GHS categories are included.

The contingency table (Table 6-26) gives an analysis of the performance based on the cut-offs shown in Figure 6-12. The data in this table indicate that the proposed cut-offs make this a very conservative model for the prediction of materials whose GHS toxicity category is greater than Category 2B with the caveat that no Category 2A materials were available for this analysis. All animal test-determined GHS Category 1 formulations were captured by this model. There were no underpredictions of Category 1 materials. In addition there were no underpredictions of Category 2B materials; all were predicted as Category 2B or higher.

What occurs as a consequence of the conservative cut-offs is that many materials are overpredicted relative to their toxicity category as determined by the animal test (LVET). Fifty percent of the Category 2B materials are overpredicted as Category 1's, and 100% of the Nonirritant materials are overpredicted as Category 2B or 1 materials

Table 6-26 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials.

LVET- Determined GHS Category	EpiOcular Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2B	NI	Total			
1	1	0	0	1	100%	NA	0%
2A	0	0	0	0	0%	0%	0%
2B	1	1	0	2	50%	50%	0%
NI	3	10	0	13	0%	100%	NA
Total	5	11	0	16	13%		
Predictivity	20%	9%	0%				
Category under predicted	NA	0%	0%				
Category over predicted	80%	91%	NA				

The discordant results for the EpiOcular assay and GHS toxicity categories without oxidizers are shown in Table 6-27. Since the prediction model did not change from the previous analysis, there was no change for the other formulation types from the analysis in Table 6-25.

Table 6-27 Prediction results for the EO assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants
Under predicted	0	0
Correctly Predicted	1 (25%)	1 (8%)
Over Predicted	3 (75%)	11 (92%)

The practical advantage of such a model is that the very low irritating materials (Category 2B and Nonirritants) can be identified and an appropriate toxicity category applied. This will clearly result in some over labeling (77% of animal-determined Nonirritants will be over labeled as Category 2B and 23% as Category 1), but the participating companies have accepted this degree of over labeling will occur. Alternatively, all of the EO predicted Category 1 materials could be retested in the BCOP assay.

<u>EPA Labeling Categories (Draize-determined)</u>

The above discussion of EPA and GHS toxicity categories (as determined by the LVET assay) utilizes a relatively small data set. However, additional EO data were available from company participants which were paired with Draize-determined EPA and GHS categories. The distribution of product categories for the additional data points is shown in Table 6-28.

Table 6-28 Distribution of product categories originally submitted with both animal eye irritation data (Draize) and EpiOcular data.

Product Categories	Number of products tested
Oxidizers	4
Surfactants	5
Solvents	6
Bases	11
Acids	1
Other	2
Total	29

Figure 6-13 presents the additional data identified by their designated product categories. Since the distribution pattern seemed to be similar to what was seen earlier, the same cut-off values as were suggested by the previous analysis of the LVET-determined EPA Categories were applied to this data set.

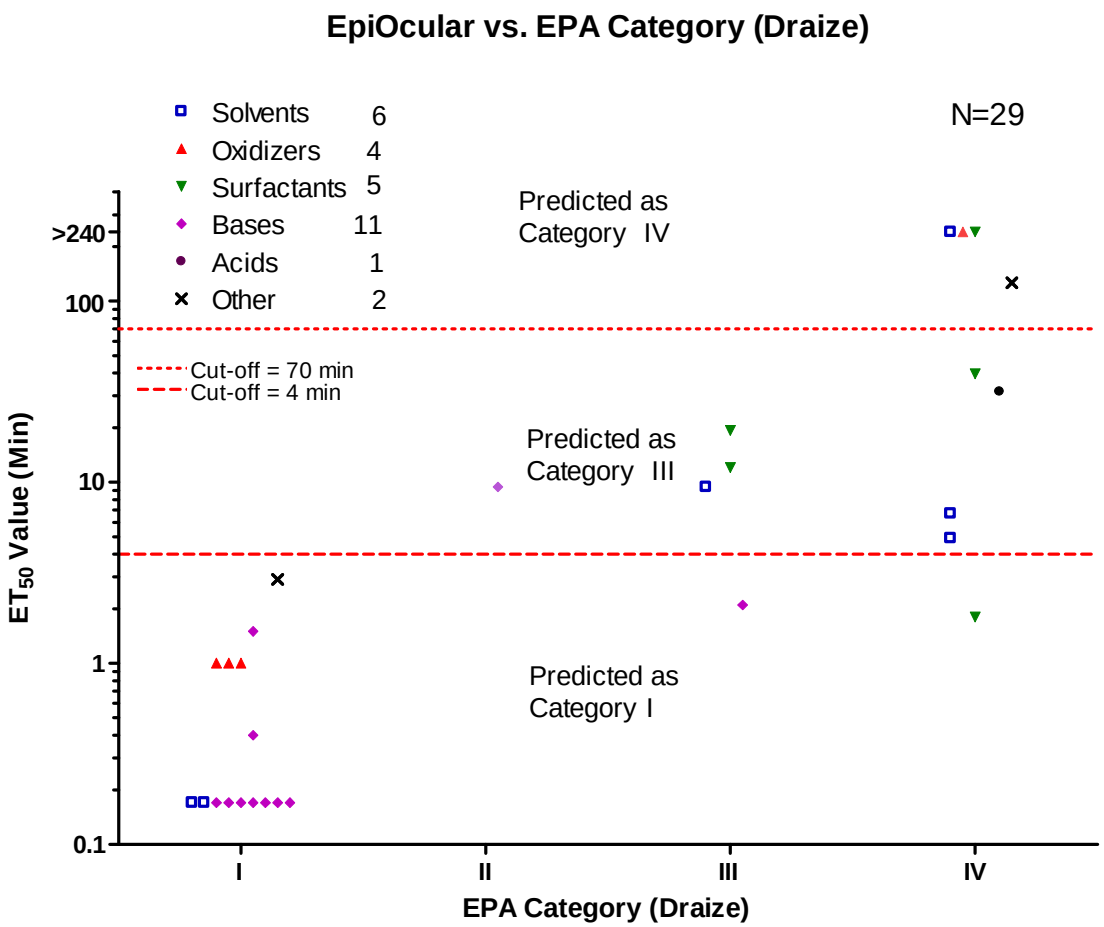


Figure 6-13 EpiOcular ET_{50} values plotted against EPA categories determined by the Draize test. Suggested cut-off values with their predicted EPA categories are included.

Again a contingency table was generated to quantitate over and underpredictions. This is shown as Table 6-29. The data in this table indicate that the proposed cut-offs are slightly less conservative than that shown with LVET-designated EPA categories. The Category I materials are correctly predicted, but the single Category II material is underpredicted. One of the Category III materials is overpredicted, but the remainder of the Category III materials are appropriately identified. Forty-four percent of the Category IV materials are overpredicted as III's and 11% (1 material) are overpredicted as I's.

Table 6-29 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the Draize test) using cut-off values of $ET_{50} \geq 70$ min = IV, and $ET_{50} < 4$ min = I. The model does not propose to identify EPA Category II materials.

Draize- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	15	0	0	15	100%	NA	0%
II	0	1	0	1	0%	0%	100%
III	1	3	0	4	75%	25%	0%
IV	1	4	4	9	44%	56%	NA
Total	17	8	4	29	76%		
Predictivity	88%	38%	100%				
Category under predicted	NA	12%	0%				
Category over predicted	12%	50%	NA				

The discordant results for the EpiOcular assay and the Draize-derived EPA toxicity categories are shown in Table 6-30. There was one underprediction of the EPA category for a base material; however, 33% of solvents, 40% of surfactants, 9% of bases, and 100% of acids were overpredicted.

Table 6-30 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	0	1 (9%)	0	0
Correctly Predicted	4 (67%)	3 (60%)	4 (100%)	9 (82%)	0	2 (100%)
Over Predicted	2 (33%)	2 (40%)	0	1 (9%)	1 (100%)	0

In this case (a different set of formulations; the Draize test used to determine EPA hazard categories) the oxidizing formulations appear to have been correctly predicted by the EO assay with the proposed cut-offs mentioned earlier. However, to parallel the analysis of the preceding section, the oxidizing formulations were removed and the data set re-evaluated in Figure 6-14.

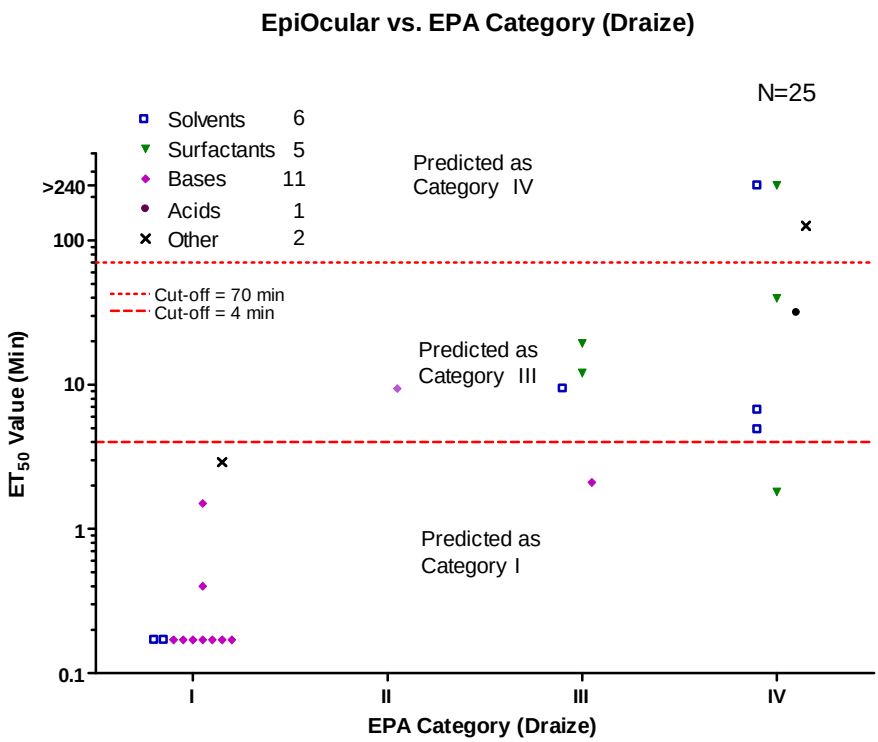


Figure 6-14 EpiOcular ET_{50} values plotted against EPA categories determined by the Draize test. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted EPA categories are included.

Again a contingency table was generated to quantitate over and underpredictions. This is shown as Table 6-31. The data in this table indicate that the proposed cut-offs are slightly less conservative than that shown with LVET-designated EPA categories. The Category I materials are correctly predicted, but the single Category II material is underpredicted. One of the Category III materials is overpredicted, but the remainder of the Category III materials are appropriately identified. Fifty percent of the Category IV materials are overpredicted as III's and thirteen percent are overpredicted as I's.

Table 6-31 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for EPA toxicity categories (determined by the Draize test) using cut-off values of ET_{50} ≥ 70 min = IV, and ET_{50} < 4 min = I. The model does not propose to identify EPA Category II materials.

Draize- Determined EPA Category	EpiOcular Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	III	IV	Total			
I	12	0	0	12	100%	NA	0%
II	0	1	0	1	0%	0%	100%
III	1	3	0	4	75%	25%	0%
IV	1	4	3	8	38%	63%	NA
Total	14	8	3	25	72%		
Predictivity	86%	38%	100%				
Category under predicted	NA	12%	0%				
Category over predicted	14%	50%	NA				

The discordant results by formulation type for the EpiOcular assay and Draize-derived EPA toxicity categories are shown in Table 6-32. Since the prediction model did not change, the results for all formulation types other than oxidizers did not change from the analysis shown in Table 6-30.

Table 6-32 Prediction results for the EO assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Bases	Acids	Other
Under predicted	0	0	1 (9%)	0	0
Correctly Predicted	4 (67%)	3 (60%)	9 (82%)	0	2 (100%)
Over Predicted	2 (33%)	2 (40%)	1 (9%)	1 (100%)	0

GHS Labeling Categories (Draize-determined)

The same data set of 29 additional materials discussed above was also evaluated for the prediction of GHS categories, however only 28 materials had sufficient animal data to provide a GHS classification. Figure 6-15 shows the distribution of the materials with respect to GHS category and EpiOcular ET_{50} value.

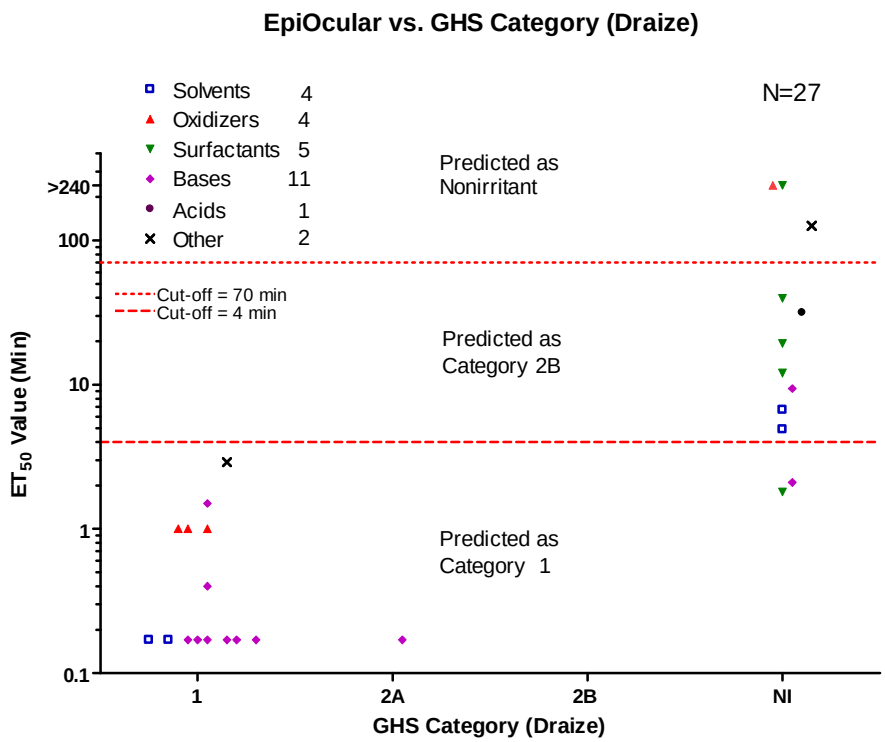


Figure 6-15 EpiOcular ET_{50} values plotted against GHS categories determined by the Draize. Suggested cut-off values with their predicted GHS categories are included.

A contingency table was generated to quantitate over and underpredictions for the GHS labeling. This is shown as Table 6-35. The data in this table indicate that the proposed cut-offs are slightly less conservative than that shown with LVET-designated GHS categories. The Category 1 materials are correctly predicted, but the single Category 2A material is overpredicted. Fify-eight percent of the Nonirritant materials are overpredicted as Category 2B.

Table 6-33 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} ≥ 70 min = NL and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials.

Draize- Determined GHS Category	EpiOcular Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2B	NI	Total			
1	14	0	0	14	100%	NA	0%
2A	1	0	0	1	0%	100%	0%
2B	0	0	0	0	NA	NA	NA
NI	2	7	3	12	25%	75%	NA
Total	17	7	3	27	63%		
Predictivity	82%	0%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	18%	100%	NA				

The discordant results for the EpiOcular assay and Draize-derived GHS toxicity categories are shown in Table 6-34. There were no underpredictions of the GHS category for any material; however, 50% of solvents, 80% of surfactants, 27% of bases, and 100% of acids were overpredicted.

Table 6-34 Discordant results for the EpiOcular assay and GHS toxicity categories.

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	0	0	0	0
Correctly Predicted	2 (50%)	1 (25%)	4 (100%)	8 (73%)	0	2 (100%)
Over Predicted	2 (50%)	4 (75%)	0	3 (27%)	1 (100%)	0

The oxidizing formulations appear to have been correctly predicted by the EO assay with the proposed cut-offs mentioned earlier. However, the oxidizing formulations were removed and the data set re-evaluated in Figure 6-16.

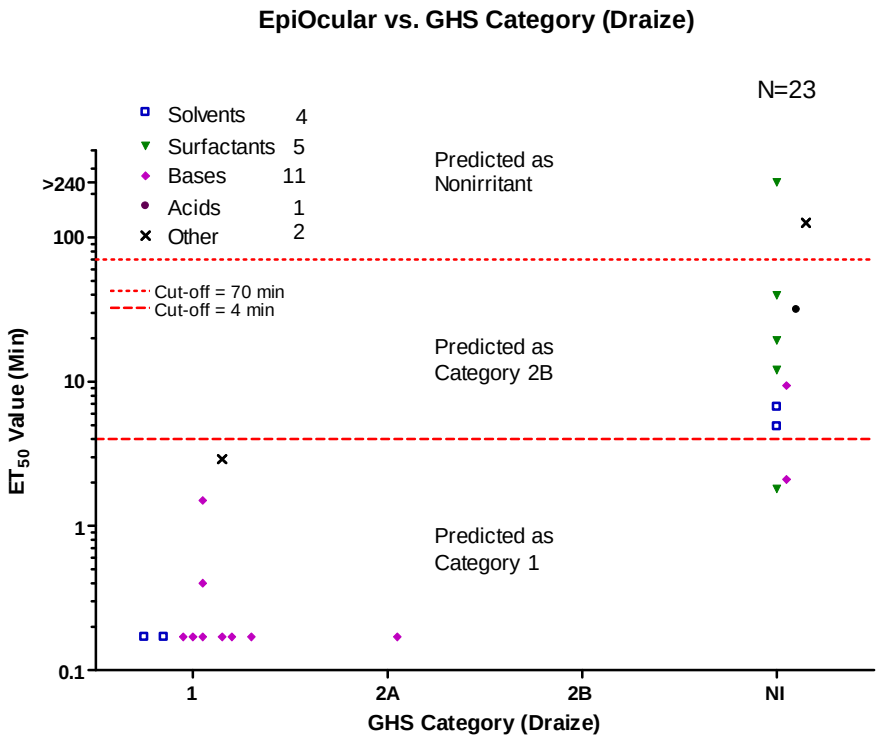


Figure 6-16 EpiOcular ET_{50} values plotted against GHS categories determined by the Draize. Oxidizers have been removed since they will be tested only in the BCOP assay. Suggested cut-off values with their predicted GHS categories are included.

A contingency table was generated to quantitate over and underpredictions for the Draize-derived GHS labeling. This is shown as Table 6-35. The data in this table indicate that the proposed cut-offs are slightly less conservative than that shown with LVET-designated GHS categories. The Category 1 materials are correctly predicted, but the single Category 2A material is overpredicted. Sixty-four percent of the Nonirritant materials are overpredicted as Category 2B.

Table 6-35 Contingency table depicting the accuracy and predictivity of the EpiOcular assay for GHS toxicity categories (determined by the LVET) using cut-off values of ET_{50} > 70 min = NI and ET_{50} < 4 min =1. The model does not propose to identify GHS Category 2A materials.

Draize- Determined GHS Category	EpiOcular Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2B	NI	Total			
1	11	0	0	11	100%	NA	0%
2A	1	0	0	1	0%	100%	100%
2B	0	0	0	0	NA	NA	NA
NI	2	7	2	11	18%	82%	NA
Total	14	7	2	23	57%		
Predictivity	79%	0%	100%				
Category under predicted	NA	0%	0%				
Category over predicted	21%	100%	NA				

The discordant results by formulation type for the EpiOcular assay and Draize-derived GHS toxicity categories are shown in Table 6-36. Since the prediction model did not change, the results for all formulation types other than oxidizers did not change from the analysis shown in Table 6-34.

Table 6-36 Discordant results for the EpiOcular assay and GHS toxicity categories.

	Solvents	Surfactants	Bases	Acids	Other
Under predicted	0	0	0	0	0
Correctly Predicted	2 (50%)	1 (25%)	8 (73%)	0	2 (100%)
Over Predicted	2 (50%)	4 (75%)	3 (27%)	1 (100%)	0

6.2.2 Conclusion for EpiOcular studies

Products used in the analysis of the EO performance had either Draize-derived or LVET-derived EPA and GHS toxicity categories. The performance of the EO assay varied somewhat depending on which of the *in vivo* assays was used, however this difference may also have been due to a different distribution of products. None the less the prediction model was determined to be the same regardless of the *in vivo* assay type. Thus the following summary is based on the combination of results from both *in vivo* assays.

The EpiOcular assay (as has been suggested by several reports (Stern, Klausner et al. 1998; Jones, Budynsky et al. 2001)) appears to be most useful at the less irritating portion of the toxicity spectrum. It is capable of identifying both EPA Category III and IV materials, although most Category IV materials will be overpredicted as Category III materials. <u>Only one of the forty-one materials (2%) was under predicted for EPA toxicity categories</u>. Over predictions were much more frequent. The corporate participants have agreed that this outcome of over labeling some materials is acceptable to them.

Very similar results were found with the prediction model for GHS categories. There were no under predictions of GHS toxicity categories regardless of whether the categories were Draize-determined or LVET-determined.

Another assay (we propose BCOP) will be used as a second tier test to differentiate EPA Category I from Category II and less irritating materials, if needed.

Historical knowledge of the performance of the EpiOcular assay plus the preceding analysis of the EpiOcular data in this BRD have led us to the following recommendations:

1) **Anti-microbial cleaning products having an oxidizing chemistry should not be tested with the EpiOcular assay.**

2) **Both water soluble and water insoluble anti-microbial cleaning products can be tested with the EpiOcular assay.**

3) **If the anti-microbial cleaning product has an ET_{50} score of <4 minutes, it is classified as EPA Category I or GHS Category 1.**

4) **If the anti-microbial cleaning product has an ET_{50} score of ≥4 minutes, but <70 minutes, it is classified as EPA Category III or GHS Category 2B.**

5) **If the anti-microbial cleaning product has an ET_{50} score of ≥70 minutes, it is classified as EPA Category IV or GHS Category NI.**

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA I or GHS 1 is actually an EPA II or a GHS 2A, it should be further tested in the BCOP assay.**

6.3 BCOP predictive capacity

6.3.1 Overview

As opposed to the Cytosensor and EpiOcular assays which were conducted almost entirely in a retrospective fashion, i.e. both the *in vitro* and *in vivo* data had been determined before the beginning of this project, the analysis of the BCOP assay proceeded sequentially with *in vitro* data being produced prospectively throughout the analysis period. An initial set of paired animal and BCOP data was used to set potential cut-off values for the various EPA categories. Subsequently new materials were received from many of the participants under code and these materials were tested in a blind fashion under GLP-compliant conditions at IIVS. Many of these materials had additional histopathological analysis which was conducted either at IIVS or at a subcontractor who had been trained in histopathological analysis of bovine corneas. The histopathological analysis was conducted while the materials were still under code.

6.3.2 Analysis using only BCOP *in vitro* scores (no histopathology)

6.3.2.1 Original company data submissions

Table 6-37 gives the distribution of materials in Figure 6-17. As seen with the analysis of the Cytosensor and the EpiOcular assays, the distribution of product categories is relatively uneven. Surfactants appear to be under represented when compared to the information available for the other two *in vitro* assays.

Table 6-37 Distribution of product categories originally submitted with both animal eye irritation data and BCOP *in vitro* data.

Product Categories	Number of products tested
Oxidizers	8
Surfactants	1
Acids	0
Bases	10
Solvents	9
Total	28

Figure 6-17 shows the distribution of the initial 28 BCOP *in vitro* scores plotted against EPA labeling categories (determined by the Draize test). Six of these materials were tested in a modified Draize protocol with a reduced volume (0.03 ml), but since the results were scored as Category I even though a reduced volume was used, it was decided that it was valid to use these data in the analysis.

It is apparent from Figure 6-17 that the distribution of BCOP *in vitro* scores across the EPA labeling categories is not random. Most EPA Category I materials

have *in vitro* scores higher than 100 and none fall below 40. In contrast all EPA Category IV materials have scores below 35. EPA Category III materials are widely spread between *in vitro* scores of 20 and 80. This distribution allowed us to set cut-off values by eye for predicting EPA labeling categories. We attempted to choose cut-offs conservatively with a bias towards having as few under predictions as was reasonable. No statistical methods were employed to construct the proposed prediction model.

It appears that all but one of the Category 1 materials is identified with a cut-off greater than an *in vitro* score of 75. It appears that it is not possible to differentiate between III's and IV's with a cut-off value, but both III's and IV's might be identified with a cut-off of below an *in vitro* score of 35. Since the BCOP assay does not differentiate between materials in the mild irritancy range as well as the other assays in this BRD, a second assay such as EO or CM may be used to demonstrate an EPA Category IV. Figure 6-17 shows a plot of BCOP *in vitro* scores versus EPA category classifications (as determined by the Draize test) with the above proposed cut-off values added.

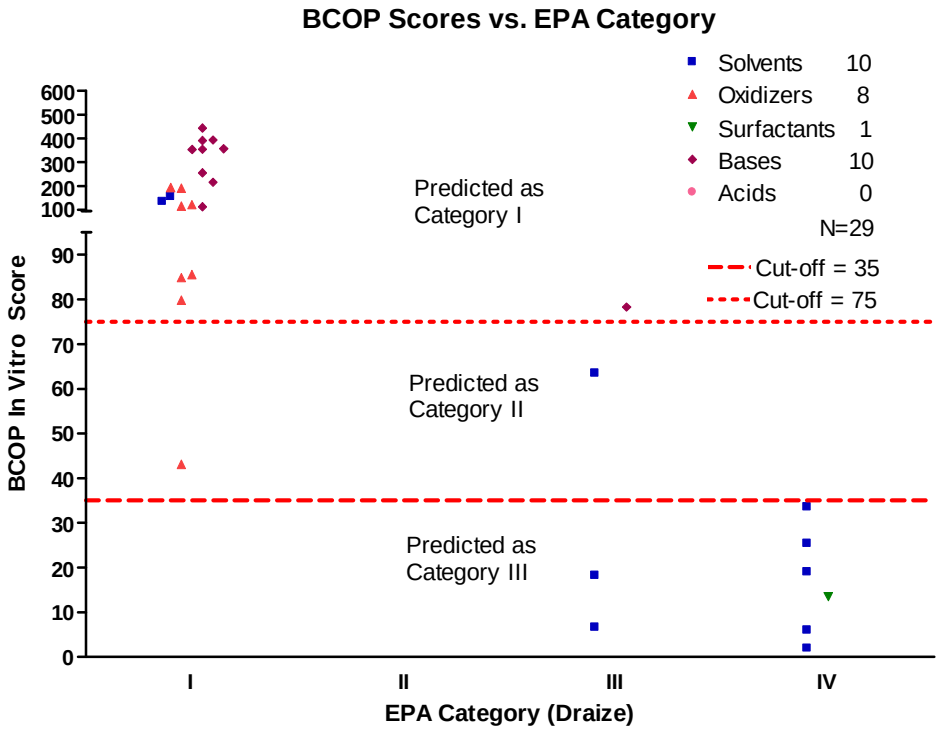


Figure 6-17 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included.

The following contingency table (Table 6-38) gives an analysis of the performance based on the preliminary cut-offs shown in Figure 6-17. The data in this table indicate that the proposed cut-offs make this a conservative model for the prediction of materials whose EPA Category is I. Ninety-five percent of the *in vivo*-

determined EPA Category I materials were captured by this model. There was only one underprediction of a Category I material. However, there were no EPA Category II materials in this initial data set; therefore, it is impossible to determine how well the predicted cut-offs actually predict Category II materials.

What occurs as a consequence of the conservative cut-offs is that many less irritating materials are overpredicted relative to their toxicity category as determined by the Draize animal test. With this prediction model all of the Category IV materials are overpredicted as Category III's, and 50% of the Category III materials are overpredicted as Category II's or Category I's.

Table 6-38 Contingency table (based on Figure 6-17) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 35 = II, and BCOP *in vitro* score < 35 = III. Although the model does propose to identify EPA Category II materials, there are no Category II's in the data set to test the hypothesis. The model does not propose to identify Category IV materials.

Draize- Determined EPA Category	BCOP Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	II	III	Total			
I	18	1	0	19	94.7%	NA	5.3%
II	0	0	0	0	0%	0%	0%
III	1	1	2	4	50%	50%	0%
IV	0	0	6	6	0%	100%	NA
Total	19	2	8	29	69%		
Predictivity	94.7%	0%	25.0%				
Category under predicted	NA	50%	0%				
Category over predicted	5.3%	50%	75.0%				

The discordant results assessed by product formulation for the BCOP assay and EPA toxicity categories are shown in Table 6-39. There was one underprediction of the EPA category for oxidizing materials; however, 60% of solvents, 100% of surfactants, and 10% of bases were overpredicted.

Table 6-39 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids
Under predicted	0	0	1 (12%)	0	0
Correctly Predicted	4 (40%)	0	7 (88%)	9 (90%)	0
Over Predicted	6 (60%)	1 (100%)	0	1 (10%)	0

The practical advantage of such a model is that the very irritating materials (Category I's) can be easily identified and an appropriate toxicity category applied. This will clearly result in some over labeling (all of *in vivo* EPA Category IV's would be over labeled as III's), unless a second tier test was used to differentiate the

Category IV's from the Category III's. The participating companies have accepted that this degree of over labeling of Category IV's will occur. As stated before, the EPA appears to concur with this type of approach since the EPA label Review Manual (2003) states (for primary eye irritation of Category IV) that "…the registrant may choose to use Category III labeling."

6.3.2.2 Further analysis

6.3.2.2.1 Additional materials tested and analyzed by EPA category

In order to test the validity of the proposed BCOP prediction model, additional cleaning products were solicited from the companies participating in creating this submission. The greatest emphasis was placed on obtaining Category II materials since none were present in the first set of submissions. However, obtaining additional compounds to test proved difficult since many of the formulations for which full animal data were available were no longer being marketed and thus would only be available if the material were reformulated specifically for this project. Additionally, obtaining EPA Category II formulations was problematic because few Category II cleaning products appear to be currently marketed (personal experience of submitter who searched retail stores). The highly aggressive Category I materials are common since they are often highly concentrated industrial and institutional cleaning products. Consumer products, on the other hand, are generally sold in a more dilute form and are less irritating (Categories III and IV) than the industrial and institutional products (personal communication, manufacturers participating in this project).

Thirty-seven additional materials (only 36 had sufficient data to obtain GHS hazard categories) for which *in vivo* rabbit data (Draize or LVET) already existed were eventually submitted during the course of this project for testing in the BCOP assay. Each of these materials was submitted in a coded form so that the laboratory conducting the BCOP assay (IIVS) would not be aware of the EPA classification already assigned to the product by the animal test.

When the 37 new materials (including more surfactants, as we had needed) and their BCOP scores were added to the database, a new plot was constructed of the BCOP *in vitro* scores versus the EPA categories (Figure 6-18). It appeared that an *in vitro* score of 75 was still a satisfactory cut-off to separate EPA Category I materials from Category II materials. **Thus the additional of the 37 new data points verified the originally postulated prediction model – an important outcome when trying to validate a prediction model.** With the addition of five EPA Category II materials it appeared that the cut-off for conservatively separating Category II from Category III materials should be lowered to an *in vitro* score of 25. This allows three of the five Category II materials to be correctly identified, as well as ensures that the three low-scoring Category I materials would not be underpredicted by more than one toxicity category.

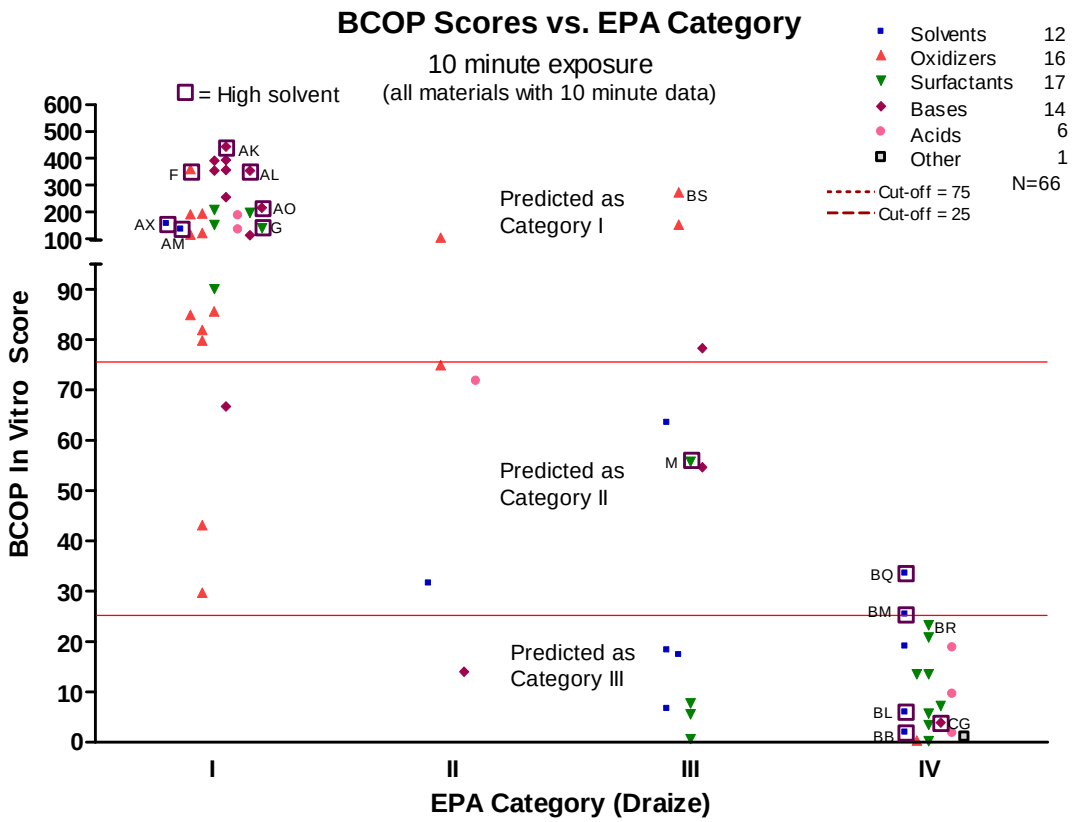


Figure 6-18 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA toxicity categories for test materials BR and BS were determined by using the results of an LVET assay. The discussion of the materials labeled as "High solvent" occurs later in this chapter.

A contingency table (Table 6-40) was constructed using the information from Figure 6-18. The results show that the BCOP assay performs well at identifying Category I materials (positive predictive value of 87.1%) while also having high sensitivity (90%) for Category I materials.

Table 6-40 Contingency table (based on Figure 6-18) depicting the accuracy and predictivity of the BCOP assay for EPA classification (determined by the Draize test) using cut-off values of *in vitro* score > 75 = I, 75 > BCOP *in vitro* score > 25 = II, and BCOP *in vitro* score < 25 = III. The model does not propose to identify Category IV materials.

Draize- Determined EPA Category	BCOP Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	II	III	Total			
I	27	3	0	30	90%	NA	10%
II	1	3	1	5	60%	20%	20%
III	3	3	6	12	50%	50%	0%
IV	0	2	17	19	0%	100%	NA
Total	31	11	24	66	54.5%		
Predictivity	87.1%	27.3%	25%				
Category under predicted	NA	27.3%	4%				
Category over predicted	12.9%	45.5%	71%				

The discordant results assessed by product formulation for the BCOP assay and EPA toxicity categories are shown in Table 6-41. There were five underpredictions of the EPA category for base and oxidizing materials; however, 50% of solvents, 53% of surfactants, 19% of oxidizers, 21% of bases, 50% of acids, and 100% of other materials were overpredicted.

Table 6-41 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	3 (19%)	2 (14%)	0	0
Correctly Predicted	6 (50%)	8 (47%)	10 (62%)	9 (64%)	3 (50%)	0
Over Predicted	6 (50%)	9 (53%)	3 (19%)	3 (21%)	3 (50%)	1 (100%)

6.3.2.2.2 Additional materials tested and analyzed by GHS toxicity category

When the 36 new materials (only 37 had sufficient raw data to calculate GHS hazard classifications) and their BCOP scores were added to the database, a new plot was constructed of the BCOP *in vitro* scores versus the GHS categories (Figure 6-19). It appeared that an *in vitro* score of 75 was a satisfactory cut-off to separate GHS Category 1 materials from GHS Category 2A materials. With the addition of six GHS Category 2A materials it appeared that the cut-off for conservatively separating GHS Category 2A from Category 2B materials should be an *in vitro* score of 25, identical to the EPA toxicity categories II/III cut-off. This allows four out of eight GHS Category 2A materials to be correctly identified, as well as ensures that the three low-scoring GHS Category 1 materials would not be underpredicted by more than one toxicity category.

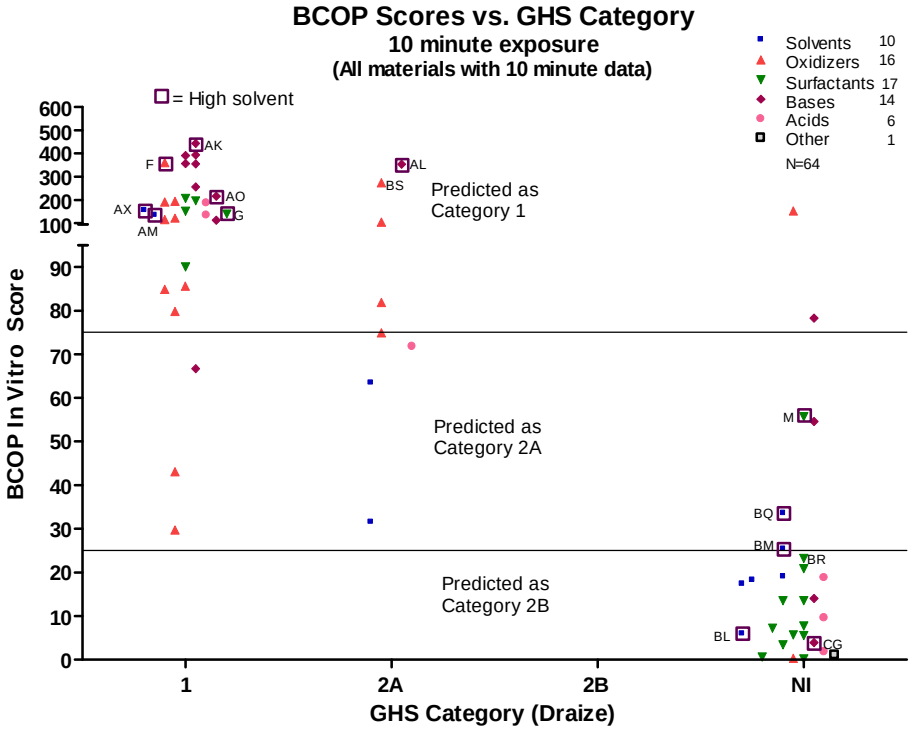


Figure 6-19 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories for test materials BR and BS were determined by using the results of an LVET assay. The discussion of the materials labeled as "High solvent" occurs later in this chapter.

A contingency table (Table 6-42) was constructed using the information from Figure 6-19. The results show that the BCOP assay performs well at identifying EPA Category 1 materials (positive predictive value of (81%) while also having high sensitivity (89%) for Category I materials.

Table 6-42 Contingency table (based on Figure 6-19) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Category NL materials.

Draize- Determined GHS Category	BCOP Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2A	2B	Total			
1	25	3	0	28	89.3%	NA	10.7%
2A	4	4	0	8	50%	50%	0%
2B	0	0	0	0	0%	0%	0%
NI	2	4	22	28	0%	100%	NA
Total	31	11	22	64	45.3%		
Predictivity	80.6%	36.4%	0%				
Category under predicted	NA	27.3%	0%				
Category over predicted	19.4%	36.4%	100%				

The discordant results assessed by product formulation for the BCOP assay and GHS toxicity categories are shown in Table 6-43. There were three underpredictions of the GHS category for base and oxidizing materials; however, 60% of solvents, 71% of surfactants, 31% of oxidizers, 36% of bases, 54% of acids, and 100% of other materials were overpredicted.

Table 6-43 Prediction results for the BCOP assay and GHS toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	2 (13%)	1 (7%)	0	0
Correctly Predicted	4 (40%)	5 (29%)	9 (56%)	8 (57%)	3 (50%)	0
Over Predicted	6 (60%)	12 (71%)	5 (31%)	5 (36%)	3 (50%)	1 (100%)

6.3.2.2.3 Analysis of anti-microbial cleaning formulations with high solvent concentrations

In the analysis presented in Figures 6-18 & 6-19, we noticed that several formulations classified as being based on a solvent chemistry for cleaning were overpredicted relative to their Draize-based classification. This phenomena of some solvents being overpredicted has been observed before and was mentioned in the conclusions of the ICCVAM/NICEATM review of the BCOP assay which states in Section 6.2.1 that:"The accuracy analysis indicated that alcohols are often overpredicted (50% to 56% [7/14 to 9/16] false positive rate depending on the classification system used) in the BCOP test method."

The formulations that are solvent-based generally contain glycol ethers or ethanol as the solvent. We examined the formulation list (see Annex B) for all the test materials which were listed as containing some amount of either "solvent" or "glycol ether", no matter what the percentage or whether they were actually categorized as "solvent" by the submitter. Thirty-one such materials were identified. These materials were then identified on the scatter plots of BCOP scores versus Draize categories, and it was found that three of these materials were overpredicted (one by one category, two by two categories). A further analysis showed that these three materials all contained either "solvent" or glycol ethers at a concentration >5%. In total, there were 13 materials that had solvent concentrations above 5%. We gave these 13 materials a new designation of "High Solvent". The identities of the High Solvent materials are shown in Figures 6-18 and 6-19 by red boxes.

Because of earlier indications that some solvent-containing materials might be overpredicted, IIVS – for the last several years – has tested such materials in the BCOP assay using two different exposure times: 3 minutes and 10 minutes. We have generally noticed that the three minute exposure gives a better prediction of the actual irritancy potential than does the 10 minute exposure. Eight of the thirteen "High Solvents" had three minute exposure data, and when we graphed these values we found that all three of the overpredicted formulations were now correctly predicted (Figure 6-20). Five high solvent materials which had been correctly

predicted as EPA Category I materials (all had BCOP scores between 157.3 and 444.3) could not be included since no three minute data had been collected when these materials were originally tested. None of the five of the materials were still available from the submitter and it was deemed too difficult to reformulate them.

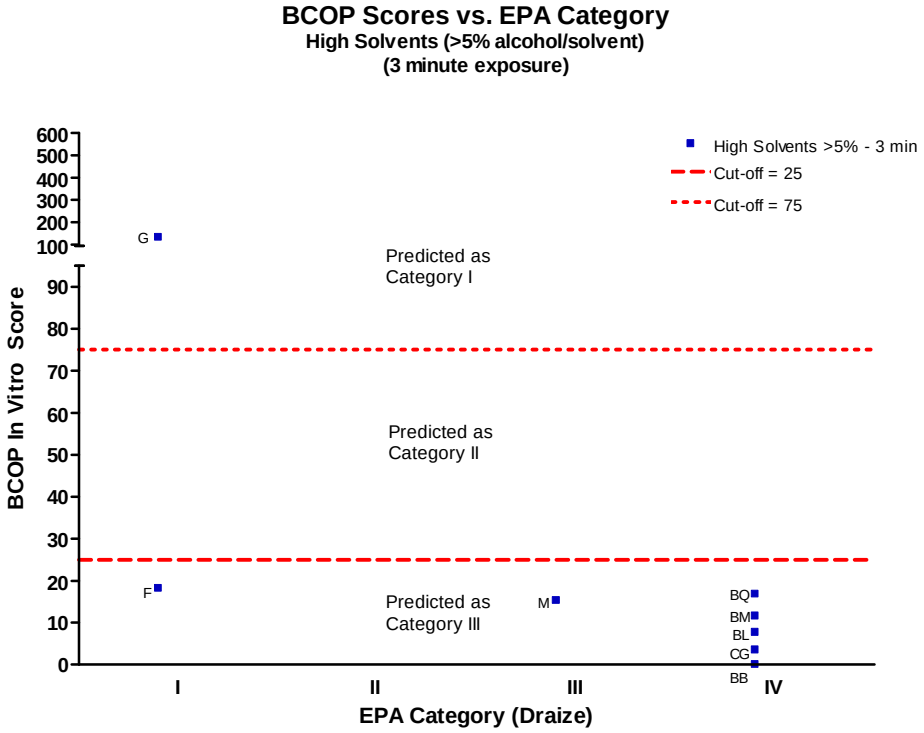


Figure 6-20 BCOP *in vitro* scores (3 minute exposure) for High Solvents are plotted against EPA categories determined by the Draize test. Five High Solvent materials had 10 minute data only and therefore are not included in this graph. Proposed cut-off values with their predicted EPA categories are included.

The remaining non-high solvent materials were then graphed as before using their 10 minute exposure time values (Figure 6-21). There are only 55 data points on Figure 6-21 because the thirteen high solvent materials are excluded.

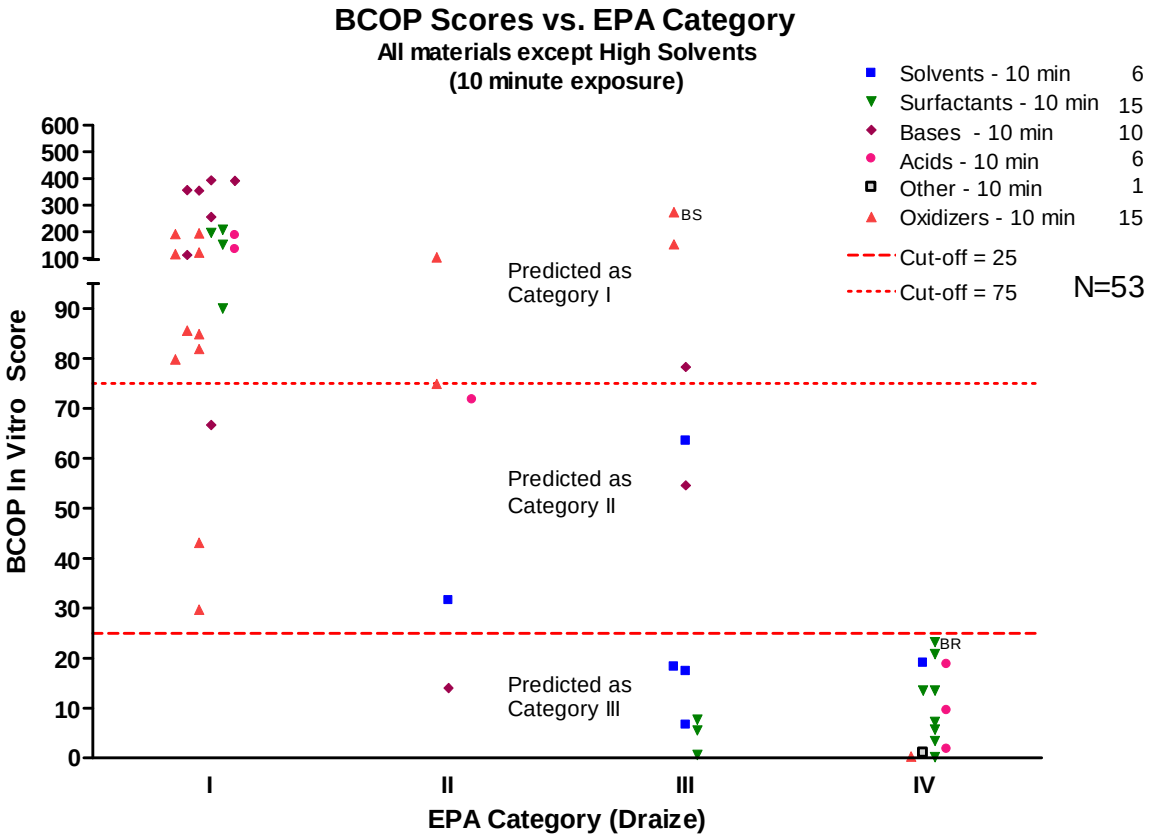


Figure 6-21 BCOP *in vitro* scores for non-High Solvent materials plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA categories for test materials BR and BS were determined by using the results of an LVET assay.

A contingency table (Table 6-44) was then created for the EPA categorization by combining the results of Figures 6-20 & 6-21. The results from this analysis are reasonably similar to that of Table 6-40 where all of the materials were recorded using their 10 minute exposure values. Using the High Solvent approach the positive predictivity for Categories I, II and III were 84%, 38% and 25%, respectively; while using the prior approach the predictivites for these categories were 87.1%, 27.3% and 28%, respectively. Thus there was some gain in the predictivity of Category II materials. However, percentages of underpredicted Category I materials increased from 9.7% to 16%, primarily because one Category I material (High Solvent) was misidentified as a Category III and five previously correctly predicted (using ten minute data) Category I High Solvent materials could not be used in the analysis since they had no 3 minute exposure data.

Table 6-44 Contingency table (based on a combination of the results from Figure 6-20 & 6-21) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 25 = II, and a BCOP *in vitro* score < 25 = III. The model does not propose to identify Category IV materials.

Draize- Determined EPA Category	BCOP Predicted EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	II	III	Total			
I	21	3	1	25	84%	NA	16%
II	1	3	1	5	60%	20%	20%
III	3	2	7	12	58%	42%	0%
IV	0	0	19	19	0%	100%	NA
Total	25	8	28	63	49%		
Predictivity	84%	38%	25%				
Category under predicted	NA	38%	7%				
Category over predicted	16%	25%	68%				

The discordant results assessed by product formulation for the BCOP assay and EPA toxicity categories are shown in Table 6-45. There were five underpredictions of the EPA category for base, oxidizing, and high solvent materials; however, 33% of solvents, 53% of surfactants, 27% of oxidizers, 20% of bases, 50% of acids, and 100% of other materials. Sixty-three percent of high solvents were overpredicted, but only by a single toxicity category.

Table 6-45 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other	High Solvents
Under predicted	0	0	2 (13%)	2 (20%)	0	0	1 (12%)
Correctly Predicted	4 (67%)	8 (50%)	9 (60%)	6 (60%)	4 (57%)	0	2 (25%)
Over Predicted	2 (33%)	8 (50%)	4 (27%)	2 (20%)	3 (43%)	1 (100%)	5 (63%)

The same analysis of using 3 minute data for the High Solvent materials was conducted using GHS categories. Figure 6-22 shows the results using the High Solvents, and Figure 6-23 shows the results with the rest of the materials. Again three previously correctly predicted High Solvent Category I materials could not be used since there was no three minute exposure data for them.

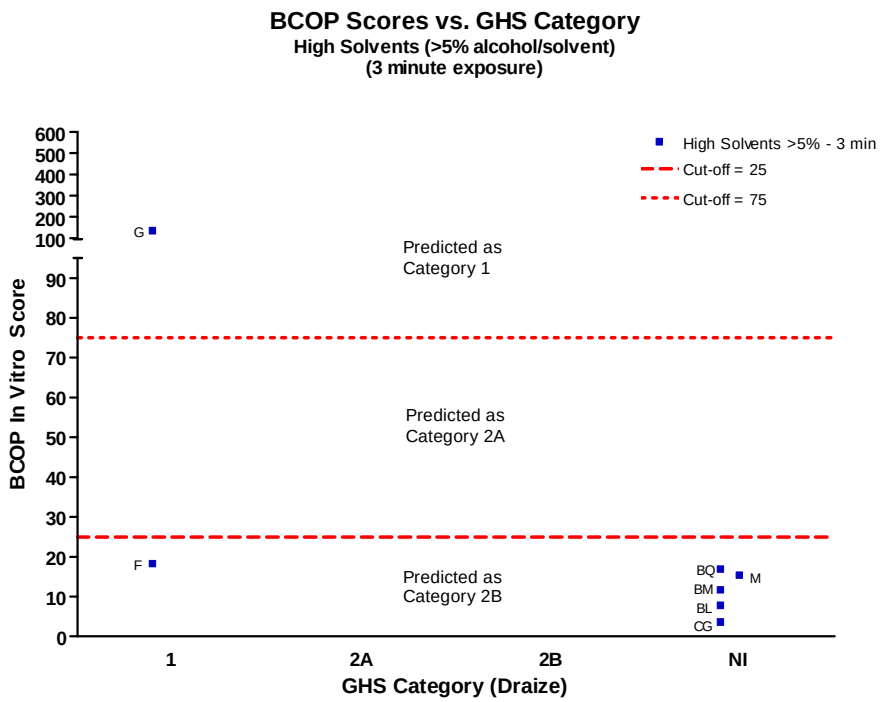


Figure 6-22 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Five materials had only 10 minute data and therefore are not included on this graph. Proposed cut-off values with their predicted GHS categories are included. Test material BB is not included due to the study criteria not being met for the GHS category.

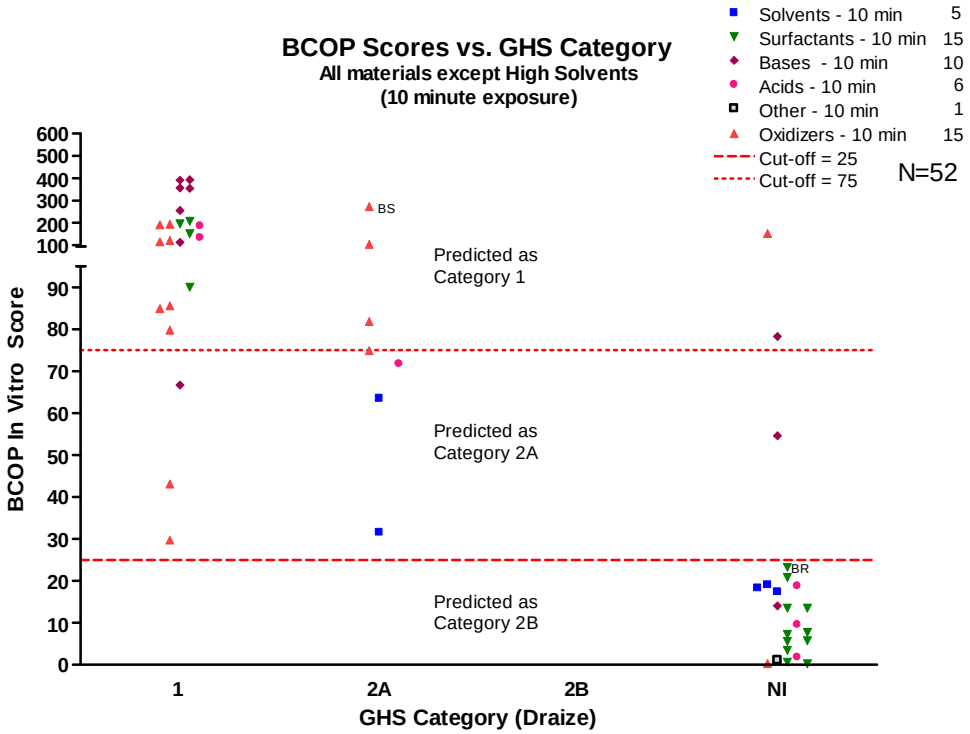


Figure 6-23 BCOP *in vitro* scores plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay.

Table 6-46 shows the results of a contingency analysis of the GHS conducted by combining the results from both graphs. As can be seen by comparing with the previous GHS category analysis in Table 6-42, the predictivity improved slightly from the original analysis, but the underprediction of Category 1 materials increased slightly from 11% to 17%, primarily because one Category 1 material (High Solvent) was misidentified as a Category 2B and five previously correctly predicted (using ten minute data) Category 1 High Solvent materials could not be used in the analysis since they had no 3 minute exposure data.

Table 6-46 Contingency table (based on Figure 6-22 & 6-23) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Category NL materials.

Draize- Determined GHS Category	BCOP Predicted GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2A	2B	Total			
1	20	3	1	24	83%	NA	17%
2A	3	4	0	7	57%	43%	0%
2B	0	0	0	0	0%	0%	0%
NI	2	1	25	28	0%	100%	NA
Total	25	8	26	59	41%		
Predictivity	80%	50%	0%				
Category under predicted	NA	38%	4%				
Category over predicted	20%	13%	96%				

The discordant results assessed by product formulation for the BCOP assay and GHS toxicity categories are shown in Table 6-47. There were four underpredictions of the EPA category for base, oxidizing, and high solvent materials; however, 60% of solvents, 73% of surfactants, 36% of oxidizers, 43% of bases, 50% of acids, 100% of other materials, and 71% of high solvents were overpredicted.

Table 6-47 Prediction results for the BCOP assay and EPA toxicity categories by product formulation type. Number of each product tested and percentage (in parentheses).

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other	High Solvents
Under predicted	0	0	2 (13%)	1 (10%)	0	0	1 (14%)
Correctly Predicted	2 (40%)	4 (27%)	8 (53%)	6 (60%)	3 (50%)	0	1 (14%)
Over Predicted	3 (60%)	11 (73%)	5 (33%)	3 (30%)	3 (50%)	1 (100%)	5 (71%)

6.3.3 Histopathology Analysis

We have previously reported (Curren, Evans et al. 2000) that certain materials, especially those with oxidizing chemistry, may be under estimated when relying only on the *in vitro* score. Often these materials cause cellular changes in the

cornea that are not manifested *in vitro* as damage by the conventional measures of opacity and permeability. Presumably similar changes *in vivo* do result in visible changes to the eye as a result of secondary recruitment and resulting migration of inflammatory cells into the corneal stroma. Thus we decided to do additional analysis of the predictive capacity of the BCOP assay by adding data from parallel studies of the histopathology of the treated corneas.

Additional rationale for the use of histopathology comes from the pioneering work of Drs. James Jester and Jim Maurer ((Maurer, Parker et al. 2002) who have shown that the area and depth of the initial ocular injury is a major predictor of the final lesion and its potential for recovery. A more complete description of this hypothesis and its relationship to the BCOP assay can be found in Annex G (Draft BCOP Histopathology Guidance Document).

For seventeen of the antimicrobial cleaning product materials, the treated bovine corneas were fixed, sectioned and stained for histopathological examination. The corneas were evaluated either by the staff of IIVS or by a subcontractor (a Board-certified veterinary pathologist) trained in histological analysis of bovine corneas. A detailed description of the types of lesions observed can be found in Annex G.

Histology was evaluated and described for the: 1) Upper, middle and lower epithelium; 2) Upper, middle, and lower stroma; and 3) Endothelium. Table 6-50 relates the histological damage observed in a cornea to a specific EPA or GHS category. Decisions as to the category assigned were based primarily on the depth of injury. For the epithelium, this was measured primarily by tissue loss or the presence of necrotic cells. For the stroma, damage was characterized by 1) abnormal chromatin condensation or vacuolated nuclei in the keratocytes, 2) significant increase in collagen matrix vacuolization, or 3) loss of keratocytes. Damage to the endothelium was evidenced by loss of cells or increased cellular vacuolization.

Table 6-48 Scoring chart for histologically apparent damage and proposed EPA and GHS toxicity category.

Extent of Damage	Suggested EPA Category	Suggested GHS Category
Cell loss or damage extending no further than midway through the epithelium.	IV	NL
Cellular damage or collagen matrix damage extending no further than the upper third of the stroma	III	2B
Cellular damage or collagen matrix damage extending no further than two-thirds of the way through the stroma	II	2A
Cellular damage or collagen matrix damage extending into the lower third of the stroma and/or causing damage to the endothelial cells.	I	1

Table 6-49 identifies the materials used to treat the corneas, the BCOP *in vitro* score, the histology results, and the final determination of the EPA toxicity category.

Table 6-49 Integration of histopathology results with BCOP *in vitro* scores to give final EPA toxicity category classification (based on prediction model of Figure 6-18). Test material code letters appear in Figure 6-24 & 6-25.

Test Material	EPA Category by Draize Test	BCOP *In vitro* Score 10 min/3 min	EPA Category by Prediction Model	Histology Results 10 min/3 min	EPA Category (Based on Histopathology) 10 min/3 min	Final EPA Category 10 min/3 min
1 (V)	IV	20.8	III	Damage observed mid-stroma	II	II
2 (I)	III	0.6	III	Upper epithelium lost	IV	III
3 (H)	II	9.2	III	Damage into lower third of stroma	I	I
4 (F) High Solvent	I	514/18.2	I/III	Damage into lower third of stroma/Damage into lower third of stroma	I/I	I/I

5 (C)	I	29.7	II	Damage into lower third of stroma	I	I
6 (X)	I	81.9	I	Damage to lower stroma	I	I
7 (Y)	II	74.9	II	Damage to lower stroma	I	I
8 (Z)	II	31.6	II	Damage to Upper Half of stroma	II	II
9 (AV)	I	191.8	I	Damage into lower third of stroma	I	I
10 (AW)	I	43.1	II	Damage greater than 50% depth of stroma	II	II
11 (BJ)	III	54.6	II	Damage through upper 2/3s of stroma	II	II
12 (AE)	I	66.7	II	Damage through top half of stroma	II	II
13 (CG) High Solvent	IV	3.9/3.5	III/III	Damage into upper quarter of stroma/Upper epithelium lost	III/IV	III/III
14 (N)	III	152.7	I	Damage into lower third of stroma	I	I
15 (BS)	III (LVET)	278.1	I	Damage into lower third of stroma	I	I
16 (BR)	IV	23.2	III	Damage through upper third of stroma	III	III
17(EG)	II	71.8	II	Damage into lower third of stroma	I	I

6.3.3.1 Analysis of the predictive capacity of BCOP including histological evaluation for EPA hazard classifications

Using the results of the above histological observations, a further analysis of the predictive capacity for EPA toxicity categories of the combination of BCOP *in vitro* score and histopathology was performed. The EPA toxicity categories are plotted against the *in vitro* score (using the same cut-offs as previously described) for both the High Solvents (Figure 6-24) and the remaining materials (Figure 6-25).

The materials which underwent histology analysis are circled and their final predicted toxicity category (as determined by Table 6-49) shown.

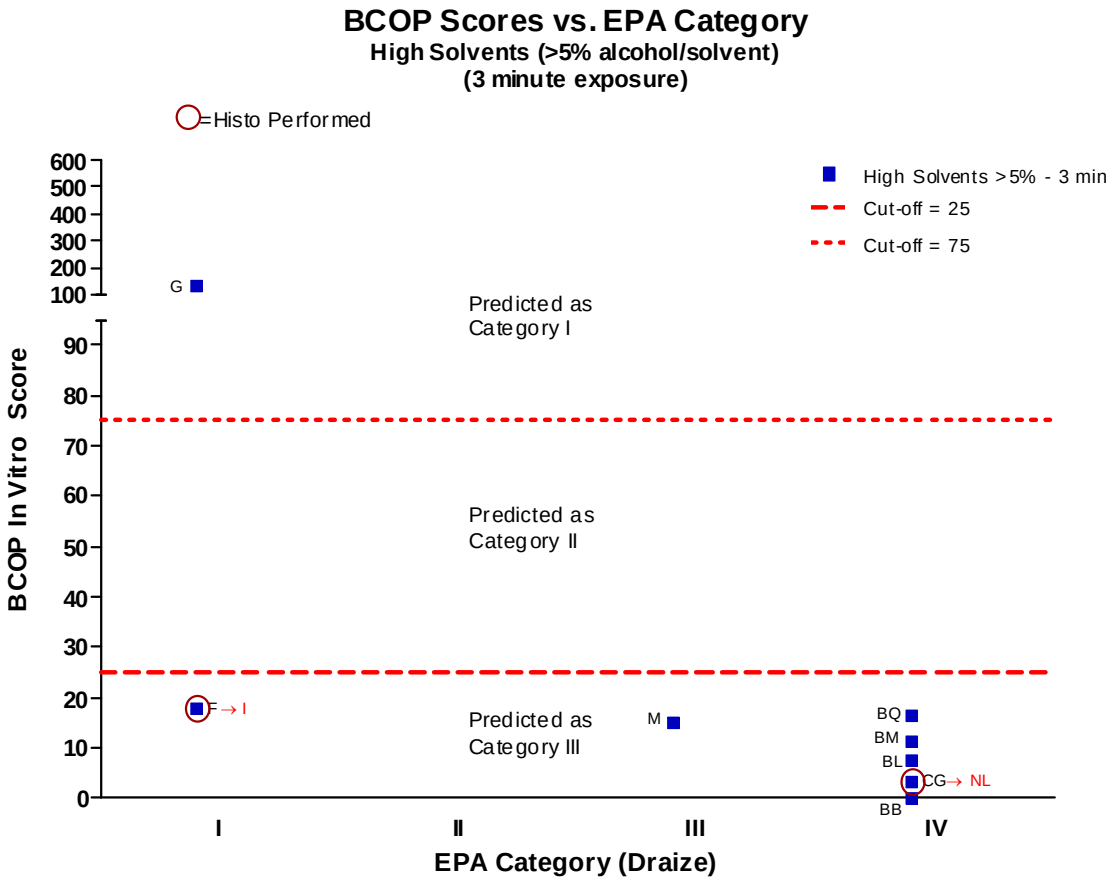


Figure 6-24 BCOP *in vitro* scores (3 minute exposure) for High Solvent formulations plotted against EPA categories determined by the Draize test. Five High Solvent materials had 10 minute data only and therefore are not included in this graph. Materials with histology-determined EPA categories are circled with the final category indicated.

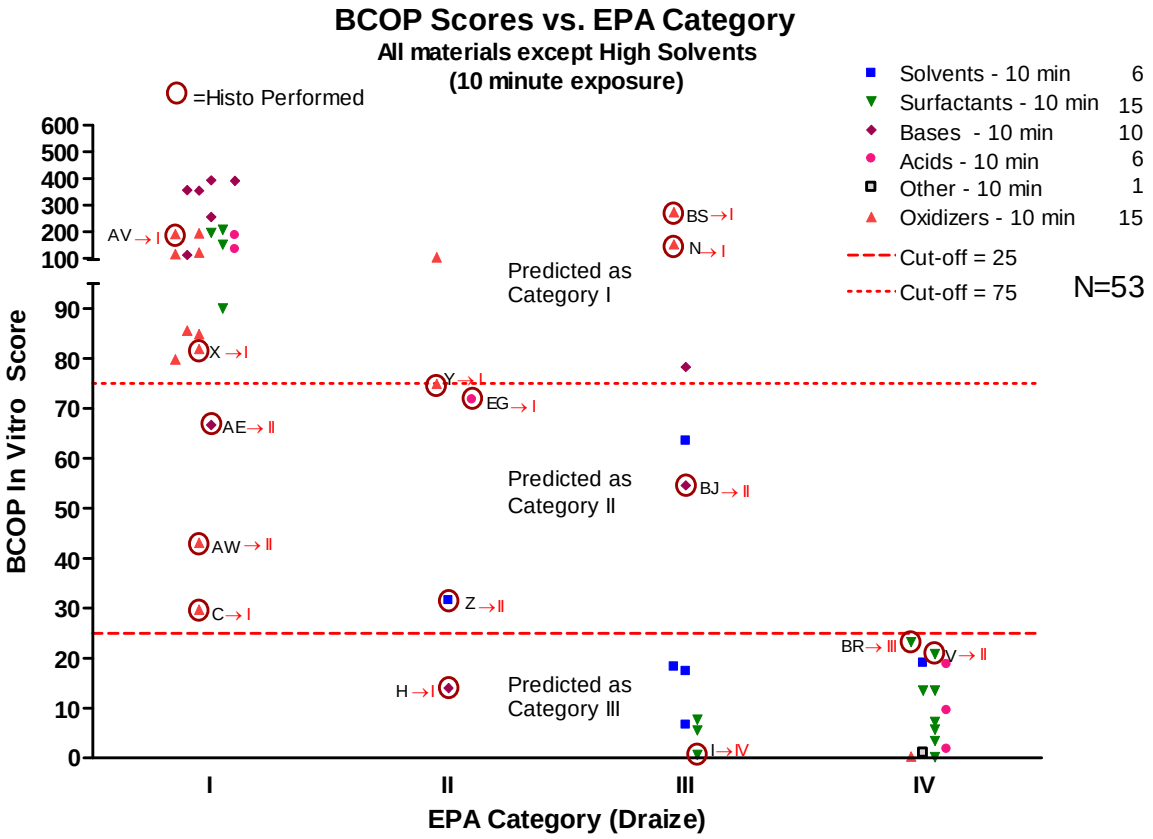


Figure 6-25 BCOP *in vitro* scores plotted against EPA categories determined by the Draize test. Proposed cut-off values with their predicted EPA categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay.

The contingency Table 6-50 shows the results of these analyses. It can be seen that adding histopathology analysis to the BCOP *in vitro* score leads to fewer EPA toxicity categories being underestimated. The sensitivity of the assay for detecting EPA category I's improves to 92% (23 of 25 Category I's identified) from 84% (Table 6-44). Similarly the underprediction of EPA Category II's improves from 20% (Table 6-44) with BCOP *in vitro* score only, to 0% when histopathology is added.

Table 6-50 Contingency table (based on Figure 6-24 & 6-25) depicting the accuracy and predictivity of the BCOP assay for EPA toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = I, 75 > BCOP *in vitro* score ≥ 25 = II, and a BCOP *in vitro* score < 25 = III, plus histopathological evaluation. The model does not propose to identify Category IV materials.

Draize- Determined EPA Category	BCOP Predicted (with histology) EPA Category				Concordance	Toxicity over predicted	Toxicity under predicted
	I	II	III	Total			
I	23	2	0	25	92%	NA	8%
II	4	1	0	5	20%	80%	0%
III	3	2	7	12	58%	42%	0%
IV	0	1	18	19	0%	100%	NA
Total	30	6	25	61	51%		
Predictivity	77%	17%	28%				
Category under predicted	NA	33%	0%				
Category over predicted	23%	50%	72%				

The discordant results assessed by type of formulation for the BCOP assay with histology and EPA toxicity categories are shown in Table 6-51. There was one underprediction each of the EPA category for bases and oxidizing materials; however, 50% of solvents, 53% of surfactants, 33% of oxidizers, 30% of bases, 67% of acids, and 100% of other materials were overpredicted.

Table 6-51 Discordant results for the BCOP assay and EPA toxicity categories.

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	1 (7%)	1 (10%)	0	0
Correctly Predicted	7 (50%)	7 (47%)	9 (60%)	6 (60%)	2 (33%)	0
Over Predicted	7 (50%)	8 (53%)	5 (33%)	3 (30%)	4 (67%)	1 (100%)

6.3.3.2 Analysis by GHS category for BCOP including histological evaluation

Using the results of the histological observations, an analysis of the predictive capacity for GHS toxicity categories of the combination of BCOP *in vitro* score and histopathology was performed. The GHS toxicity categories are plotted against the *in vitro* score (using the same cut-offs as previously described) for both the High Solvents (Figure 6-26) and the remaining materials (Figure 6-27). The materials which underwent histology analysis are circled and their final predicted toxicity category (as determined by Table 6-49) shown.

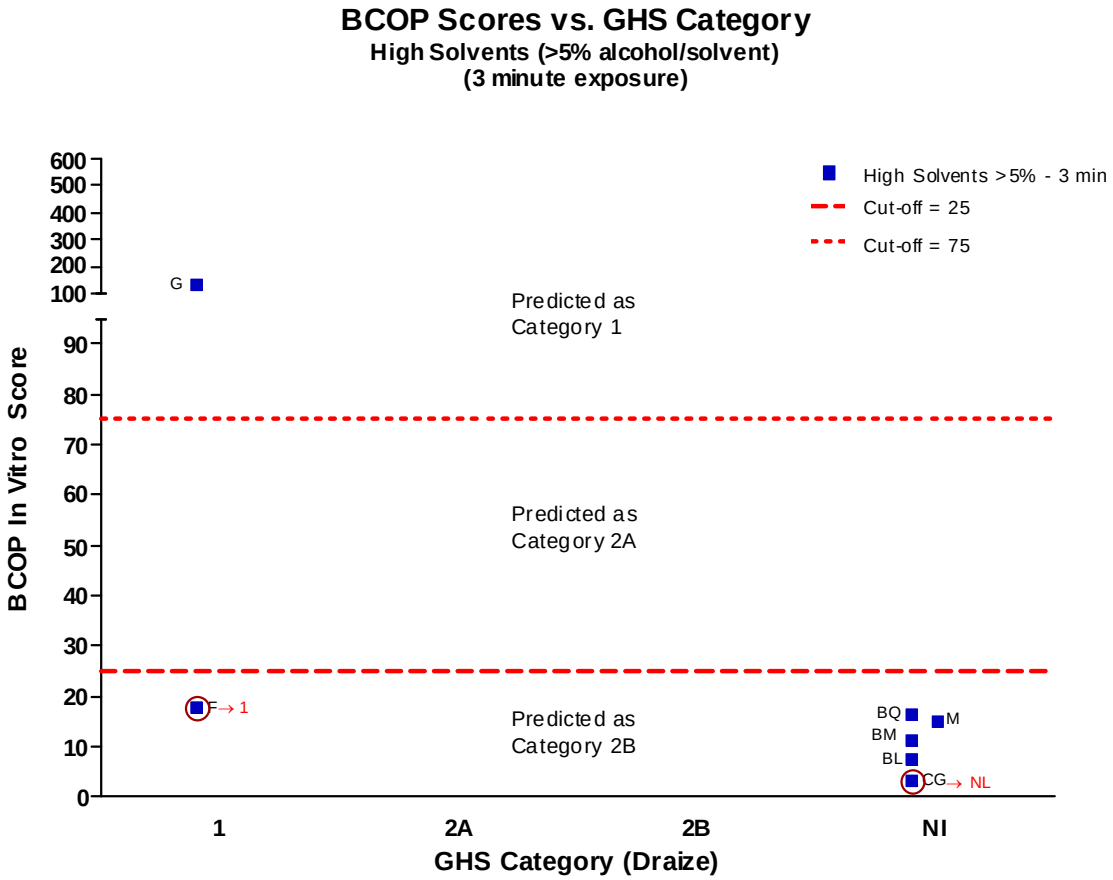


Figure 6-26 BCOP *in vitro* scores (3 minute exposure) for High Solvent formulations plotted against GHS categories determined by the Draize test. Five High Solvent materials had only 10 minute data and therefore are not included on this graph. Proposed cut-off values with their predicted GHS categories are included. Materials with histology-determined EPA categories are circled with the final category indicated. Test material BB is not included due to the study criteria not being met for the GHS category.

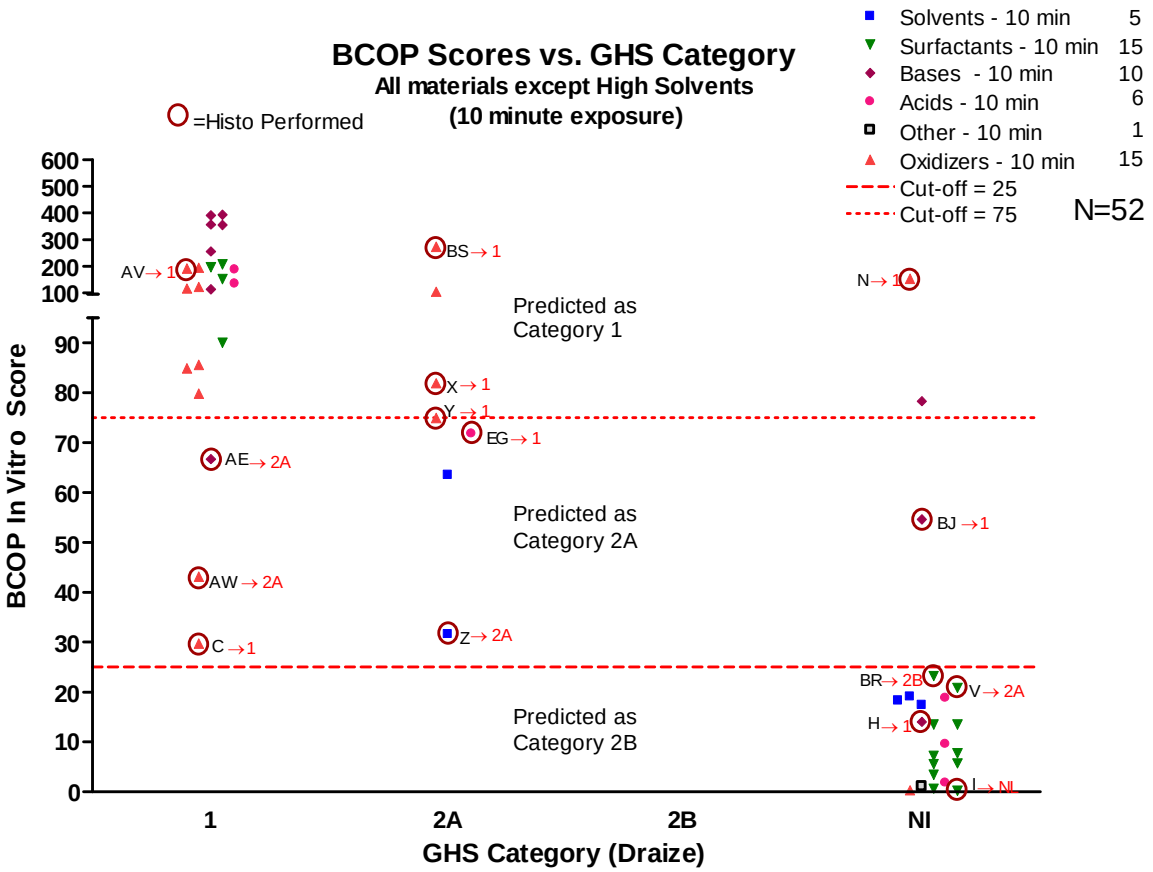


Figure 6-27 BCOP *in vitro* scores for non-High solvent materials plotted against GHS categories determined by the Draize test. Proposed cut-off values with their predicted GHS categories are included. The EPA categories of test materials BR and BS were determined using the LVET assay. Materials with histology-determined EPA categories are circled with the final category indicated.

The contingency Table 6-52 shows the results of these analyses. It can be seen that adding histopathology analysis to the BCOP *in vitro* score leads to fewer GHS toxicity categories being underestimated. The sensitivity of the assay for EPA category I's improves to 92% (22 of 24 Category 1's identified) from 83% (Table 6-46). However, the overprediction of GHS 2A materials increases from 43% (Table 6-46) to 71%.

Table 6-52 Contingency table (based on Figure 6-26 & 6-27) depicting the accuracy and predictivity of the BCOP assay for GHS toxicity categories (determined by the Draize test) using cut-off values of a BCOP *in vitro* score ≥ 75 = 1, 75 > BCOP *in vitro* score ≥ 25 = 2A, and a BCOP *in vitro* score < 25 = 2B. The model does not propose to identify Nonirritant materials.

Draize- Determined GHS Category	BCOP Predicted (with histology) GHS Category				Concordance	Toxicity over predicted	Toxicity under predicted
	1	2A	2B	Total			
1	22	2	0	24	92%	NA	8%
2A	5	2	0	7	29%	71%	0%
2B	0	0	0	0	0%	0%	0%
NI	4	1	23	28	0%	100%	NA
Total	31	5	23	59	41%		
Predictivity	71%	40%	0%				
Category under predicted	NA	40%	0%				
Category over predicted	29%	20%	100%				

The discordant results assessed by type of formulation for the BCOP assay with histology and GHS toxicity categories are shown in Table 6-53. There was one underprediction each of the GHS category for bases and oxidizing materials; however, 67% of solvents, 73% of surfactants, 40% of oxidizers, 30% of bases, 67% of acids, and 100% of other materials were overpredicted.

Table 6-53 Discordant results for the BCOP assay and GHS toxicity categories.

	Solvents	Surfactants	Oxidizers	Bases	Acids	Other
Under predicted	0	0	1 (7%)	1 (10%)	0	0
Correctly Predicted	4 (33%)	4 (27%)	8 (53%)	6 (60%)	2 (33%)	0
Over Predicted	8 (67%)	11 (73%)	6 (40%)	3 (30%)	4 (67%)	1 (100%)

6.3.3.3 Conclusions from analysis of the BCOP predictive capacity

The BCOP assay appears to be most useful at the most irritating portion of the irritation spectrum. It is capable of identifying EPA toxicity category I, II, and III materials. When histology was considered, only two of the sixty-one materials (3%) were under predicted for EPA toxicity categories, and these were only under predicted by a single toxicity category. Over predictions were much more frequent, but this was driven by the fact that the BCOP assay seems incapable of clearly differentiating between Category III and Category IV materials. In fact 64% (18 of 28) of the over predictions were the result of EPA category IV materials being predicted as EPA category III materials. Very similar results were found with the prediction model for GHS categories. Sixty-seven percent (22 of 33) of the over predictions were Non Irritating materials over predicted as category 2B's. The corporate participants have agreed that this outcome of over labeling some materials is acceptable to them.

Many High Solvent (>5% solvent) materials were overpredicted using the traditional 10 minute exposure. Predictions improved when a three minute exposure was used, and this shorter exposure is recommended for future use with high Solvent formulations.

Importantly, when BCOP testing (and selected histopathology) was conducted on 37 new antimicrobial cleaning product formulations, the results fit the pattern of the originally hypothesized prediction model. Thus the preliminary hypothesis was supported, lending considerable weight to the validity of this prediction model.

We report that histopathology can be performed on treated corneas – this allows for possible underpredictions to be discovered. Another assay (we propose Cytosensor or EpiOcular) can be used as a second tier test to differentiate EPA Category III from Category IV and less irritating materials, if needed. This will reduce the over prediction rate of the entire testing strategy.

The preceding analysis of the BCOP data has led us to the following recommendations:

1) **In general, when testing anti-microbial cleaning product formulations, the BCOP assay should be conducted with a ten minute exposure.**

2) **If the anti-microbial cleaning product contains a solvent at the level of 5% or greater, it should be tested with a three minute exposure.**

3) **All anti-microbial cleaning products having an *In Vitro* Score ≥75 should be classified as an EPA Category I or a GHS Category 1. <u>No histopathology needs to be conducted.</u>**

4) **Anti-microbial cleaning products having an *In Vitro* Score <75 and ≥ 25 are given a preliminary classification of EPA Category II or GHS Category 2A. They should be further assessed with a histopathological evaluation and given the final categorization of whichever determination (in vitro score or histological evaluation) is more severe.**

5) **Anti-microbial cleaning products having an *In Vitro* Score <25 are given a preliminary classification of EPA Category III or GHS Category 2B. They should be further assessed with a histopathological evaluation and given the final categorization of whichever determination (in vitro score or histological evaluation) is more severe.**

6) **(Optional) To determine if an anti-microbial cleaning product which was categorized as either EPA III or GHS 2B is actually an EPA IV**

or a GHS NI, it should be further tested in either the Cytosensor or EpiOcular assays.

6.4 Strategic testing approach

Data from each of the three proposed assays shows that they each have a set of strengths and weaknesses. What is especially apparent is that the Cytosensor and EpiOcular assays do not have the ability to clearly separate Category I and II materials from each other. However, both are able to identify a proportion of the very mild EPA category IV or GHS NI materials. Thus the utility of these two assays is in the mild end of the irritation spectrum.

The BCOP assay, in contrast, is able to separate the Category II materials from the Category I materials, but it is not able to differentiate between Category III materials and the Category IV materials.

The strategy we propose is a tiered testing process that can be initiated with any of the three assays (as long as the test material is physically compatible with that assay). Thus for a suspected highly aggressive material one would start with the BCOP assay. This test might immediately identify the material as a Category I or II material, in which case the testing would end. However, the BCOP assay might merely identify the material as less than a Category II (either a Category III or Category IV). If the manufacturer does not need to know whether or not it might be a Category IV, the testing again could stop with the material being labeled a Category III material. On the other hand if the manufacturer thought the material might be a Category IV, the testing could move to the next tier of EpiOcular or Cytosensor assay where the predictive capacity of the latter two assays in the mild range may appropriately identify the toxicity category.

Alternatively, testing could start with either the EpiOcular or the Cytosensor assay which have the ability to classify a material as a Category III, Category IV, or greater than a Category III. Since these assays cannot differentiate between Category II and Category I materials, the material would have to be labeled Category I if no more testing was desired. However, the material could be tested in the next tier (BCOP assay) if it was important to the manufacturer to know whether the material might be a Category II.

The second test in this tiered approach is always an option. If the exact irritation category is not required, and the manufacturer can accept an over prediction for a material whose *in vitro* score was ambiguous, then a single test can always provide the necessary information.

7 Test Method Reliability

Test method reliability will be assessed by reporting on the intralaboratory repeatability I (runs conducted in a single laboratory within a short period of time [days]), intralaboratory reproducibility II (runs conducted in a single laboratory within an extended period of time [months]) and interlaboratory reproducibility (between-laboratory repeatability). Typically the reliability of a method is assessed utilizing the data sets contained within the BRD. However, in this submission there are insufficient examples of repeated studies to provide a rigorous assessment of reproducibility for each of the methods. Therefore, information from other sources (a Background Review Document of the Cytosensor submitted to ECVAM, a Background Review Document for the EpiOcular method submitted to ECVAM and a Background Review Document on the Bovine Cornea Opacity and Permeability Test Method prepared by NICEATM which has been reviewed by an Expert Panel) will be presented as evidence for reproducibility. Where data to assess reproducibility are available within this submission they will be appropriately utilized. Table 7-1 details the study, reported results, number of replicates, and format of the available data for each type of variability study described in Section 7.

The reliability of each of the three methods proposed for this testing strategy will be addressed individually below. For most of the examples, reproducibility is assessed by calculating the CV for MRD_{50}, ET_{50}, or *in vitro* score values obtained from identical materials.

Table 7-1 Description of the results reported for each variability study.

Studies	Results Reported					
	Variability	*In Vitro* Test	Number of test substances	No. of operators	No. of assays	Format of data (raw data, summary results, other)
EC/HO	Intralaboratory	Cytosensor	35	Unknown	3 - 5	Summary; Mean
COLIPA	Intralaboratory	Cytosensor	26 and 29[1]	Unknown	≥3 for each of 2 labs	Raw data; Mean, SD and CV
20 chemicals from EC/HO and the COLIPA study run by same lab	Intralaboratory	Cytosensor	16	Unknown	≥3	See above for individual studies
Microbiological Associates/IIVS positive control	Intralaboratory	Cytosensor	1	Multiple	629	Raw data; mean; SD and CV
Formulations submitted for this BRD	Intralaboratory (within run and between experiment)	EpiOcular	15	Unknown	1 – 3 exp. 2 rep./exp.	Raw data; Mean, SD and CV
MatTek/IIVS positive control	Intralaboratory	EpiOcular	1 tested multiple times	Multiple	Unknown[2]	Mean, SD and CV
Colgate-Palmolive Phase II & III	Interlaboratory	EpiOcular	19 – 4 labs 54 – 2 labs	Unknown	2 - 4	Raw data; Mean, SD and CV
Formulations submitted for this BRD	Intralaboratory (within run)	BCOP	75	Multiple	3 - 5	Raw data; Mean, SD and CV
Formulations submitted for this BRD	Intralaboratory (between experiments)	BCOP	5	Multiple	2 - 6	Mean, SD, and CV
Gautheron et al.	Interlaboratory	BCOP	51	Multiple	4-12	Mean
Balls et al.	Interlaboratory	BCOP	59	Multiple	Unknown	Mean, CV
Southee et al.	Interlaboratory	BCOP	16	Multiple	Unknown	Mean, CV

[1]One lab tested 26 materials while a second lab tested 29 materials
[2]The same material was evaluated in two labs over a nine year period. The exact number of experiments is unknown.

7.1 Cytosensor

7.1.1 Cytosensor intralaboratory repeatability I

The within-run reproducibility could not be assessed for the materials submitted for this study because the Cytosensor reports were not readily available. However, within-run reproducibility has been demonstrated for the Cytosensor in several large international validation studies as presented in a BRD submitted to ECVAM. Table 7-2 presents the results extracted from the BRD for the EC/HO validation study (Balls, Botham et al. 1995) for a group of 32 materials (a mixture of surfactant and non-surfactant materials). For this study the mean CV was 38.9% and the median CV was 30.5%. The distribution of product categories for the within-laboratory reproducibility of the CM is shown in Table 7-3.

Table 7-2 Within-laboratory reproducibility of CM from archived data that was originally obtained at Microbiological Associates, Inc. for the EC/HO study (Balls, Botham et al. 1995). The protocol utilized the CM using Transwells and an 810 second exposure time. At least triplicate runs were performed.

Chemical	Formulation Type	Values used to calculate the reported Mean MRD_{50} (mg/mL)					Values not used to calculate the reported Mean MRD_{50} (mg/mL)					Results using the values used to calculate the reported mean				Results using all values other than > or < values to calculate mean[a]			
		Assay 1	Assay 2	Assay 3	Assay 4	Assay 5	Range Finding Assay	Assay A	Assay B	Assay C	Assay D	Average MRD_{50} (mg/mL)	N	SD	CV (%)	Average MRD_{50} (mg/mL)	N	SD	CV (%)
2,5-Dimethylohexanediol	SO		151.36	151.36	165.96		>5.00					156.22	3	8.43	5.4	156.22	3	8.43	5.4
Acetone	SO		144.54	165.96	114.82		>5.00					141.77	3	25.68	18.1	141.77	3	25.68	18.1
Ammonium nitrate	Other		169.82	125.89	144.54		>5.00					146.75	3	22.05	15.0	146.75	3	22.05	15.0
Benzalkonium chloride	SU			0.48	0.43	0.50	0.35	>3.50	<3.00			0.47	3	0.04	8.2	0.44	4	0.07	15.6
Benzalkonium chloride	SU		0.95	0.98	1.38		3.01					1.10	3	0.24	21.7	1.58	4	0.97	61.5
Benzalkonium chloride [1]/[2]	SU		5.01	3.89	7.08		3.02					5.33	3	1.62	30.4	4.75	4	1.75	36.9
Cetylpyridinium bromide	SU		0.98	1.15	0.93		3.09					1.02	3	0.11	11.1	1.54	4	1.04	67.6
Cetylpyridinium bromide	SU		1.26	0.72	2.69		1.20					1.56	3	1.02	65.3	1.47	4	0.85	57.8
Cetylpyridinium bromide	SU		128.82	91.20	91.20		>5.00					103.74	3	21.72	20.9	103.74	3	21.72	20.9
Ethanol*	SO		162.18	67.61	147.91		>5.00					125.90	3	50.98	40.5	125.90	3	50.98	40.5
Ethyl acetate	SO				53.70		>4.70	>4.75	>4.75			53.70	1			53.70	1		
Gammabutyrolactone	Other		93.33	144.54	112.20		>5.00					116.69	3	25.90	22.2	116.69	3	25.90	22.2
Glycerol	SO		218.78		131.83	204.17	>5.00		>5.38			184.93	3	46.56	25.2	184.93	3	46.56	25.2
Imidazole*	SU			23.44	22.39	23.44	24.55	>5.25				23.09	3	0.61	2.6	23.45	4	0.88	3.8
Isobutanol	SO		29.51	28.84	27.54		4.60					28.63	3	1.00	3.5	22.62	4	12.04	53.2
Isopropanol	SO		89.13	93.33	91.20		4.74					91.22	3	2.10	2.3	69.60	4	43.27	62.2
L-Aspartic acid	AC		0.85	2.63	0.71		2.75					1.40	3	1.07	76.7	1.74	4	1.11	63.8
Methyl acetate	SO		89.13	93.33	93.33		4.98					91.93	3	2.43	2.6	70.19	4	43.52	62.0
Methyl ethyl ketone	SO		53.70	45.71	52.48		4.82					50.63	3	4.31	8.5	39.18	4	23.17	59.1
Potassium cyanate	Other		50.12	30.90	30.20		26.92					37.07	3	11.30	30.5	34.53	4	10.53	30.5
Promethazine HCl	Other		1.66	1.48	1.32		1.02					1.49	3	0.17	11.5	1.37	4	0.27	19.7
Pyridine	Other		28.84	30.20	30.90		21.88					29.98	3	1.05	3.5	27.96	4	4.14	14.8
Sodium hydroxide	AL		0.87	2.24	2.09		7.24					1.73	3	0.75	43.3	3.11	4	2.82	90.7
Sodium hydroxide	AL		14.13	26.92	8.32	21.88	4.47					17.81	4	8.23	46.2	15.14	4	9.29	61.4
Sodium lauryl sulfate	SU		3.02	2.88		3.24	3.09					3.05	3	0.18	5.8	3.06	4	0.15	4.8
Sodium lauryl sulfate	SU		0.81	0.54	0.49		2.95					0.61	3	0.17	28.5	1.20	4	1.18	98.3
Thiourea	Other		50.12	48.98	53.70		>5.00					50.93	3	2.47	4.8	50.93	3	2.47	4.8
Trichloroacetic acid	AC		1.82	5.13	1.62		3.09					2.86	3	1.97	69.0	2.92	4	1.61	55.3
Trichloroacetic acid	AC		15.85	9.77	16.98		3.93					14.20	3	3.88	27.3	11.63	4	6.03	51.9
Triton X-100	SU		1.35	2.57	2.19		2.95					2.04	3	0.62	30.7	2.26	4	0.69	30.3
Triton X-100 [1]/[2]	SU		4.57	2.82	3.02		3.16					3.47	3	0.96	27.6	3.39	4	0.80	23.5
Tween 20	SU		7.76	4.79	4.57		7.41					5.71	3	1.78	31.3	6.13	4	1.69	27.5
[b]		not irr. at sol. (4.00)					>4.00	>4.00	>4.00	>4.00									
[b]		not irr. at sol. (4.00)					>4.00	>4.00	>4.00	>4.00									
[b]		not irr. at sol. (5.50)					>5.00	5.02	>5.50	>5.50	>5.50					5.02			
Mean															**23.9**				**38.9**
Median															**21.7**				**30.5**

[a] mean calculated using the values from assays 2-5, plus range finding assay if an actual value was obtained, i.e. < and > values were omitted.
[b] the identity of these chemicals is unknown
AC - Acid, AL - Alkaline, SO - Solvent, SU - Surfactant

Table 7-3 Distribution of product categories for the within-laboratory reproducibility of the CM.

Product Categories	Number of products tested
Surfactants	12
Acids	3
Bases	2
Solvents	9
Other	6
Unknown	3
Total	**35**

Results from a second international validation study organized by the European cosmetics trade association COLIPA are presented in Tables 7-4 through 7-7. These data come from both surfactant materials (Tables 7-4 and 7-6) and non-surfactant materials (Tables 7-5 and 7-7). Two different laboratories participated in this study and the individual results for each are presented. It can be seen that the first laboratory had a mean CV of 19.7% for the surfactant materials and a mean CV of 15.4% for the non-surfactant materials. The second laboratory had a mean CV of 14.3% for the surfactant materials and a mean CV of 10.4% for the non-surfactant materials. The distribution of product categories for the within-laboratory reproducibility of the COLIPA study is shown in Table 7-8.

For more details of each of these studies plus within-run repeatability from several additional studies please see Section 3. Within-laboratory reproducibility in the Cytosensor BRD. This can be provided to ICCVAM by the authors of this BRD after its review by ECVAM, or ICCVAM can request it directly from ECVAM at any time.

Table 7-4 Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived Microbiological Associates, Inc. data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-nine total materials were tested.

Substance	Formulation Type	MRD_{50} (mg/ml)	SD	CV (%)	Number of replicates
Shampoo #1 normal	**SU**	**0.75**	**0.21**	**28.7**	**3**
Eye make-up remover	**SU**	**87.77**	**1.17**	**1.3**	**3**
Triton X-100 1%	**SU**	**21.17**	**4.21**	**19.9**	**3**
Tween 20	**SU**	**9.50**	**5.31**	**55.9**	**3**
SLS 3%	**SU**	**3.23**	**0.65**	**20.2**	**3**
Triton X-100 5%	**SU**	**4.66**	**0.52**	**11.1**	**3**
Benzalkonium chloride 1%	**SU**	**4.11**	**0.89**	**21.6**	**3**
SLS 15%	**SU**	**0.52**	**0.02**	**3.5**	**3**
SLS 30%	**SU**	**0.31**	**0.02**	**5.8**	**3**
Triton X-100 10%	**SU**	**2.47**	**0.57**	**23.0**	**3**
Benzalkonium chloride 5%	**SU**	**0.81**	**0.10**	**12.7**	**3**
Benzalkonium chloride 10%	**SU**	**0.32**	**0.07**	**21.0**	**3**
Pump Deodorant	**SU**	**19.35**	**9.38**	**48.5**	**3**
Gel cleaner	**SU**	**5.68**	**2.37**	**41.8**	**3**
Shampoo - baby	**SU**	**2.51**	**0.96**	**38.1**	**3**
Hair styling lotion	**SU**	**164.82**	**7.98**	**4.8**	**3**
Liquid soap #1	**SU**	**0.88**	**0.03**	**3.5**	**3**
Mouthwash	**SU**	**37.84**	**3.55**	**9.4**	**3**
Skin cleaner	**SU**	**0.63**	**0.10**	**16.3**	**3**
Cetylpyridinium bromide 6%	**SU**	**1.36**	**0.20**	**14.5**	**3**
Polyethylene glycol 400	**SU**	**296.50**	**34.17**	**11.5**	**3**
Mean				19.7	
Median				16.3	

Table 7-5 Non-Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived Microbiological Associates, Inc. data created for the COLIPA study for non-surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-nine materials were tested.

Substance	Formulation Type	MRD_{50} (mg/ml)	SD	CV (%)	Number of replicates
Methyl ethyl ketone	**SO**	**54.18**	**3.16**	**5.8**	**3**
Imidazole	**SU**	**18.84**	**5.52**	**29.3**	**3**
Propylene glycol		**265.07**	**3.54**	**1.3**	**3**
Glycerol	**SO**	**214.83**	**25.35**	**11.8**	**3**
Sodium hydroxide 1%	**AL**	**9.09**	**1.00**	**11.0**	**3**
Isopropanol	**SO**	**52.59**	**17.20**	**32.7**	**3**
Sodium hydroxide 10%	**AL**	**4.33**	**0.15**	**3.5**	**3**
Trichloroacetic acid 30%	**AC**	**1.12**	**0.31**	**28.1**	**3**
Mean				15.4	
Median				11.4	

Table 7-6 Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived CellTox AB data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-six materials were tested.

Substance	Formulation Type	MRD_{50} (mg/ml)	SD	CV (%)	Number of replicates
Shampoo #1 normal	**SU**	**0.72**	**0.06**	**8.1**	**3**
Eye make-up remover	**SU**	**99.31**	**1.00**	**1.0**	**3**
Triton X-100 1%	**SU**	**16.79**	**0.73**	**4.3**	**3**
Tween 20	**SU**	**3.49**	**0.62**	**17.7**	**3**
SLS 3%	**SU**	**2.78**	**0.07**	**2.7**	**3**
Triton X-100 5%	**SU**	**2.42**	**0.07**	**2.7**	**3**
Benzalkonium chloride 1%	**SU**	**4.33**	**1.19**	**27.4**	**3**
SLS 15%	**SU**	**0.51**	**0.02**	**3.3**	**3**
Triton X-100 10%	**SU**	**1.24**	**0.28**	**22.9**	**3**
Benzalkonium chloride 5%	**SU**	**1.38**	**0.12**	**8.9**	**3**
Benzalkonium chloride 10%	**SU**	**0.31**	**0.05**	**16.4**	**3**
Pump Deodorant	**SU**	**47.74**	**28.34**	**59.4**	**3**
Gel cleaner	**SU**	**5.47**	**1.20**	**22.0**	**3**
Shampoo - baby	**SU**	**2.15**	**0.73**	**33.7**	**3**
Hair styling lotion	**SU**	**292.01**	**6.07**	**2.1**	**3**
Liquid soap #1	**SU**	**0.68**	**0.10**	**14.0**	**3**
Mouthwash	**SU**	**46.85**	**9.20**	**19.6**	**3**
Skin cleaner	**SU**	**0.76**	**0.05**	**6.0**	**3**
Polyethylene glycol 400	**SU**	**316.23**	**0.00**	**0.0**	**3**
Mean				14.3	
Median				8.9	

Table 7-7 Non-Surfactant Materials – COLIPA Within-laboratory reproducibility of CM from archived CellTox AB data created for the COLIPA study for surfactant materials (Brantom, Bruner et al. 1997; Harbell, Osborne et al. 1999). The protocol utilized L929 cells and an 810 second exposure. Twenty-six materials were tested.

Substance	Formulation Type	MRD_{50} (mg/ml)	SD	CV (%)	Number of replicates
Imidazole	**SU**	**26.03**	**0.99**	**3.8**	**3**
Propylene glycol		**218.86**	**7.59**	**3.5**	**3**
Glycerol	**SO**	**208.70**	**3.06**	**1.5**	**3**
Isopropanol	**SO**	**124.51**	**25.26**	**20.3**	**3**
Sodium hydroxide 1%	**AL**	**13.59**	**5.11**	**37.6**	**3**
Sodium hydroxide 10%	**AL**	**0.60**	**0.01**	**1.9**	**3**
Trichloroacetic acid 30%	**AC**	**1.24**	**0.05**	**4.2**	**3**
Mean				10.4	
Median				3.8	

Table 7-8 Distribution of product categories for the within-laboratory reproducibility of the COLIPA study

Product Categories	Number of products tested
Surfactants	21
Acids	1
Bases	2
Solvents	3
Other	
Total	**27**

7.1.2 Cytosensor intralaboratory reproducibility II

There were no examples of intralaboratory reproducibility for studies submitted specifically for this BRD. However, there is one example illustrated in the Cytosensor BRD which was submitted to ECVAM. This example was created by observing that some identical materials were tested in both the EC/HO validation study for eye irritation and the COLIPA-sponsored validation study which occurred approximately 21 months later. Although the materials are listed by the same name in each study, it is unclear whether the materials were actually identical (as far as purity and the presence of contaminants) since they were procured at a different time and possibly from different sources.

Table 7-9 presents the results for 11 surfactant materials tested by one laboratory during the EC/HO study and the COLIPA study. Both cetylpyridinium bromide (10%) and polyethylene glycol 400 were deemed incompatible with the test system in one study, but not in the other. They are both listed to show that there is some variability associated with determining whether or not a material is compatible with the test apparatus. Similarly for the non-surfactant materials shown in Table 7-10, ethyl acetate was considered incompatible with the Cytosensor in one study but not in the other. It can be seen that mean CV for the 9 surfactant material tested in both studies was 17.4%, and the mean CV for the 7 non-surfactant materials tested in both studies was 32.5%. The distribution of product categories for the

intralaboratory reproducibility of the COLIPA and EC/HO comparison is shown in Table 7-11.

Table 7-9 Surfactant materials - Comparison of the MRD_{50} values for testing conducted approximately 21 months apart

Substance	Formulation Type	COLIPA Mean MRD_{50} (mg/mL) [CV%] MA	EC/HO Mean MRD_{50} (mg/mL) [CV%] SM 31	Mean MRD_{50} (mg/mL)	SD	CV (%)
Tween 20	SU	9.50 [55.9]	5.53 [31.3]	7.50	2.83	37.7
Sodium lauryl sulphate 3%	SU	3.23 [20.2]	3.04 [6.0]	3.13	0.15	4.8
Triton X-100 5%	SU	4.66 [11.1]	3.39 [27.6]	4.03	0.90	22.3
Benzalkonium chloride 1%	SU	4.11 [21.6]	5.16 [30.4]	4.62	0.72	15.6
Sodium lauryl sulphate 15%	SU	0.517 [3.5]	0.60 [28.5]	0.56	0.06	10.9
Triton X-100 10%	SU	2.47 [23.0]	1.96 [30.7]	2.21	0.37	16.6
Benzalkonium chloride 5%	SU	0.811 [12.7]	1.09 [21.7]	0.96	0.20	21.4
Benzalkonium chloride 10%	SU	0.321 [21.0]	0.47 [8.5]	0.39	0.10	26.3
Cetylpyridinium bromide 6%	SU	1.36 [14.5]	1.35 [65.3]	1.35	0.01	0.6
Cetylpyridinium bromide 10%	SU	*	1.02 [11.1]	*	*	*
Polyethylene glycol 400	SU	296.5 [11.5]	*	*	*	*
Mean		**[19.5]**	**[26.11]**			**17.4**
Median		**[17.4]**	**[28.05]**			**16.6**

* - Material determined to be unsuitable for testing

Table 7-10 Non-surfactant materials - Comparison of the MRD_{50} values for testing conducted approximately 21 months apart

Substance	Formulation Type	COLIPA Mean MRD_{50} (mg/mL) [CV%] MA	EC/HO Mean MRD_{50} (mg/mL) [CV%] SM 31	Mean MRD_{50} (mg/mL)	SD	CV (%)
Imidazole	SU	18.8 [29.3]	23.1 [2.7]	20.95	3.04	14.5
Glycerol	SO	214.8 [11.8]	180.7 [26.6]	197.75	24.11	12.2
Sodium hydroxide 1%	AL	9.09 [11.0]	16.2 [50.0]	12.65	5.03	39.8
Isopropanol	SO	52.6 [32.7]	91.2 [2.3]	71.90	27.29	38.0
Methyl ethyl ketone	SO	54.2 [5.8]	50.5 [8.5]	52.35	2.62	5.0
Sodium hydroxide 10%	AL	4.33 [3.5]	1.60 [43.3]	2.97	1.93	65.1
Trichloroacetic acid 30%	AC	1.12 [28.1]	2.47 [69.0]	1.80	0.95	53.2
n-Butyl acetate	SO	*	*	*	*	*
Ethyl acetate	SO	*	53.7	*	*	*
Mean		**[17.5]**	**[28.9]**			**32.5**
Median		**[11.8]**	**[26.6]**			**22.6**

* - Material determined to be unsuitable for testing

Table 7-11 Distribution of product categories for the intralaboratory reproducibility of the CM.

Product Categories	Number of products tested
Surfactants	11
Acids	1
Bases	2
Solvents	5
Other	0
Total	**19**

Another data set that can give information about intralaboratory reproducibility is the response of a single material over time. The following description is extracted from the Cytosensor BRD submitted to ECVAM.

"The CM instrument was first used by the *in vitro* toxicology staff at Microbiological Associates, Inc. in 1994. At that time the practice of maintaining a graphical record of the results of the positive control material – 10% SLS in sterile, deionized water – was begun (Figure 7-1). This practice has continued through the transfer of the instrument and staff to the Institute for *In Vitro* Sciences, Inc. in 1997, and continues to this day. Table 7-12 presents a summary of the results for 629 assays conducted over a 12 plus year period as well as the results from the last 94 assays conducted over the last two years. That little change has occurred in the absolute MRD_{50} in the last 12 years can be inferred from the 12 year average of 0.0799 mg/mL versus the last two year's average of 0.0775 mg/mL. The average CV calculated over the last 12 years is 14.3%. Over the last approximately 2 years the average CV has increased to 18.9%.

Table 7-12 Positive Control Data of SLS completed at IIVS

Substance	Dates	No. of Assays	Mean MRD_{50} (mg/mL)	SD	CV (%)
SLS	April, 14 1994 – June 30, 2006	629	0.0799	0.011	14.3
SLS	March 2, 2004 - June 30, 2006	94	0.0775	0.015	18.9

SLS MRD_{50} values are plotted on a control graph with upper and lower cut-off ranges graphed at two SD of all data (March 2004 – June 2006). Assays performed on days when the MRD_{50} fell outside of the two SD range (5 points on this graph) were repeated. Because on some days more than one SLS control was run, some points may overlap such that it may appear that fewer than 94 values are plotted.

It appears from these data that there is good long term with-in lab reproducibility for a single material."

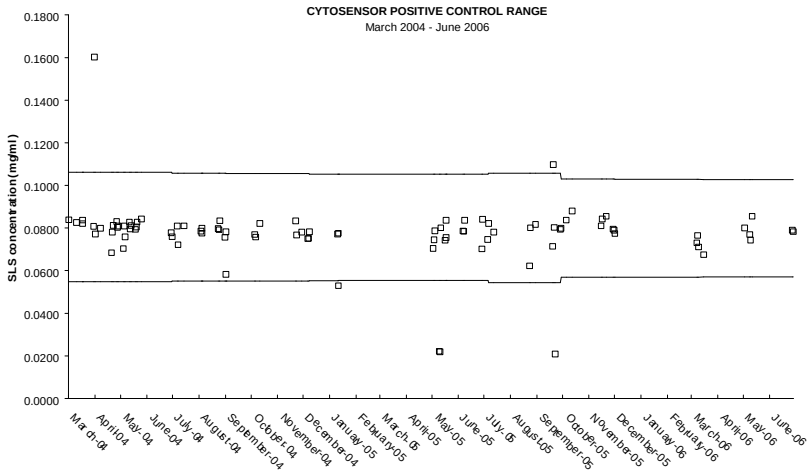


Figure 7-1 Graph of 10% SLS (positive control) MRD_{50} values obtained at IIVS over a 28-month period.

Additional information on the intralaboratory reproducibility can be found in Section 3. Within-laboratory reproducibility of the Cytosensor BRD submitted to ECVAM.

7.1.3 Cytosensor interlaboratory reproducibility

There were no examples of interlaboratory reproducibility for studies submitted specifically for this BRD. However, there are two main examples which are presented in the Cytosensor BRD submitted to ECVAM. One is from the EC/HO international validation study for eye irritation where four Cytosensor laboratories participated and the other is from the COLIPA validation study which had two Cytosensor laboratories participating.

The results from the EC/HO study are broken down into two categories: 1) surfactants (Table 7-13), and 2) non-surfactant materials (Table 7-14). For the 11 surfactants (only one laboratory found that polyethylene glycol 400 was compatible with the Cytosensor) the mean CV for the 4 laboratories was 37.0% and for the non-surfactants was 50.6%. Not all laboratories found that all the non-surfactant materials met the testing criteria, so the number of laboratories testing each of the 32 materials ranged from 1 to 4. The distribution of product categories for the interlaboratory reproducibility of the EC/HO study is shown in Table 7-15.

Table 7-13 Surfactant Materials - Between-laboratories reproducibility of CM results from EC/HO study.

Chemical	Formulation Type	Conc. tested	MRD_{50} Values (mg/mL)				Mean MRD_{50} (mg/mL)	SD	CV (%)
			CM 30	CM 31	CM 32	CM 33			
Cetylpyridinium bromide	SU	10%	0.78	1.02	2.34	0.89	1.26	0.73	58.2
Cetylpyridinium bromide	SU	6%	0.6	1.35	0.44	1.11	0.87	0.43	48.8
Benzalkonium chloride	SU	5%	1.15	1.09	0.98	1.28	1.13	0.12	11.1
Benzalkonium chloride	SU	10%	0.26	0.47	0.38	0.44	0.39	0.09	24.2
Triton X-100	SU	10%	1.61	1.96	1.50	2.22	1.82	0.33	18.0
Sodium lauryl sulfate	SU	15%	0.62	0.60	0.51	0.74	0.62	0.10	15.5
Benzalkonium chloride [1]/[2]	SU	1%	4.71	5.16	4.65	3.58	4.53	0.67	14.8
Triton X-100 [1]/[2]	SU	5%	1.90	3.39	5.09	2.53	3.23	1.39	43.0
Sodium lauryl sulfate	SU	3%	2.71	3.04	3.74	3.64	3.28	0.49	15.0
Tween 20	SU	100%	1.52	5.53	4.98	1.06	3.27	2.31	70.5
Cetylpyridinium bromide	SU	0.10%	48.19	102.33	7.76	180.30	84.65	74.62	88.1
Polyethylene glycol 400	SU	100%	*	*	*	363.92	*	*	*
Mean									**37.0**
Median									**24.2**

* Participating laboratory did not test the chemical because it determined that chemical was not compatible with the test system.

Table 7-14 Non-surfactant materials - Between-laboratories reproducibility of CM results from EC/HO study.

Chemical	Formulation Type	Conc. tested	MRD_{50} Values (mg/mL)				Mean MRD_{50} (mg/mL)	SD	CV (%)
			CM 30	CM 31	CM 32	CM 33			
Sodium hydroxide	AL	10%	2.28	1.60	2.67	2.49	2.26	0.47	20.8
Trichloroacetic acid	AC	30%	1.69	2.47	0.81	2.20	1.79	0.73	40.7
Captan 90 concentrate		100%	*	*	*	*	*	*	*
Chlorhexidine		100%	*	*	*	*	*	*	*
Cyclohexanol	SO	100%	15.49	*	0.58	*	8.03	10.5	131.3
Quinacrine		100%	*	*	1.08	*	*	*	*
Promethazine HCl		100%	1.35	1.48	0.81	1.45	1.27	0.31	24.4
Parafluoraniline		100%	*	*	3.47	*	*	*	*
Acetone	SO	100%	153.82	140.28	139.00	162.18	148.82	11.15	7.5
n-Hexanol	SO	100%	*	*	*	*	*	*	*
1-Naphthalene acetic acid		100%	12.11	*	*	*	*	*	*
Sodium oxalate		100%	*	*	*	*	*	*	*
Isobutanol	SO	100%	28.84	28.64	22.54	31.62	27.91	3.83	13.7
Imidazole	SU	100%	22.75	23.07	0.18	48.75	23.69	19.85	83.8
2-Ethyl-1-hexanol	SO	100%	*	*	*	*	*	*	*
4-Carboxybenzaldehyde		100%	*	*	*	*	*	*	*
Methyl ethyl ketone	SO	100%	55.72	50.47	78.16	47.97	58.08	13.77	23.7
Pyridine		100%	1.54	29.99	15.92	31.48	19.73	14.01	71.0
1-Naphthalene acetic acid		100%	*	*	*	*	*	*	*
2,2-Dimethylbutanoic acid	AC	100%	*	*	*	*	*	*	*
Gammabutyrolactone		100%	79.98	114.82	0.91	179.47	93.79	74.39	79.3
Thiourea		100%	50.12	50.93	*	47.97	49.68	1.53	3.1
n-Octanol	SO	100%	*	*	*	*	*	*	*
Methyl acetate	SO	100%	61.09	91.83	116.14	109.65	94.68	24.64	26.0
L-Aspartic acid	AC	100%	1.11	1.17	*	*	1.14	0.04	3.6
Benzoyl-L-tartaric acid		100%	0.81	*	*	*	*	*	*
Potassium cyanate		100%	28.18	36.06	9.40	50.82	31.11	17.25	55.4
Isopropanol	SO	100%	83.18	91.20	87.10	143.55	101.26	28.39	28.0
Sodium perborate, $4H_20$		100%	0.11	*	*	3.27	1.69	2.24	132.6
Dibenzyl phosphate	AC	100%	0.75	*	*	*	*	*	*
2,5-Dimethylohexanediol	SO	100%	75.21	155.96	6.21	156.31	98.67	72.25	73.2
Methyl cyanoacetate		100%	42.95	*	0.13	*	21.54	30.28	140.5
Sodium hydroxide	AL	1%	28.18	16.22	32.36	31.62	27.1	7.48	27.6
Ethanol	SO	100%	97.05	117.49	123.03	110.41	111.99	11.22	10.0
2,6-Dichlorobenzoyl chloride		100%	*	*	*	*	*	*	*
Ammonium nitrate		100%	40.27	145.55	27.99	*	71.27	64.62	90.7
Ethyl-2-methylacetoacetate		100%	*	*	0.40	*	*	*	*

Chemical	Formulation Type	Conc. tested	MRD_{50} Values (mg/mL)				Mean MRD_{50} (mg/mL)	SD	CV (%)
			CM 30	CM 31	CM 32	CM 33			
Ethyl acetate	SO	100%	*	53.70	*	*	*	*	*
Maneb		100%	*	*	*	*	*	*	*
Fomesafen		100%	*	*	*	*	*	*	*
Tetraaminopyrimidine sulfate		100%	1.05	*	*	*	*	*	*
Toluene		100%	*	*	*	*	*	*	*
n-Butyl acetate		100%	*	*	*	*	*	*	*
Trichloroacetic acid	AC	3%	13.90	13.80	16.29	16.11	15.03	1.36	9.0
Methyl isobutyl ketone		100%	*	*	0.81	*	*	*	*
Ethyl trimethyl acetate		100%	*	*	*	*	*	*	*
Methylcyclopentane		100%	*	*	*	*	*	*	*
Glycerol	AL	100%	121.62	180.72	8.26	208.93	129.88	88.87	68.4
Mean									**50.6**
Median									**28.0**
Mean when all four labs tested material									**39.0**

* Participating laboratory did not test the chemical because it determined that chemical was not compatible with the test system.

Table 7-15 Distribution of product categories for the interlaboratory reproducibility of the EC/HO study.

Product Categories	Number of products tested
Surfactants	12
Acids	5
Bases	3
Solvents	12
Other	
Total	

For more details of each of these studies plus additional information on interlaboratory reproducibility please see Section 5. Between-laboratory reproducibility in the Cytosensor BRD.

The results from the COLIPA study are broken down into three categories: 1) surfactants (Table 7-16), 2) surfactant-based formulations and mixtures (Table 7-17), and 3) non-surfactants, ingredients and mixtures (Table 7-18). For the 13 surfactants, both laboratories tested 10, but one laboratory found 2 materials that did not meet the testing criteria. The mean CV for the 2 laboratories was 23.3% for the surfactants, 16.5% for the 7 surfactant-based formulations and mixtures and 32.5% for the 9 non-surfactants. The distribution of product categories for the interlaboratory reproducibility of the COLIPA study is shown in Table 7-19.

Table 7-16 Surfactant Materials - Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study.

Chemical	Formulation Tested	Conc. tested	MRD_{50} Values (mg/mL)		Mean MRD_{50} (mg/mL)	SD	CV (%)
			MA	CT AB			
Triton X-100 1%	SU	1%	21.17	16.79	18.98	3.1	16.3
Tween 20	SU	100%	9.5	3.49	6.50	4.25	65.4
SLS 3%	SU	3%	3.23	2.78	3.00	0.32	10.6
Triton X-100 5%	SU	5%	4.66	2.42	3.54	1.58	44.7
Benzalkonium chloride 1%	SU	5%	4.11	4.33	4.22	0.16	3.7
SLS 15%	SU	1%	0.52	0.51	0.51	0.01	1.0
SLS 30%	SU	100%	0.31	*	*		
Triton X-100 10%	SU	15%	2.47	1.24	1.85	0.87	46.8
Benzalkonium chloride 5%	SU	30%	0.81	1.38	1.1	0.4	36.7
Benzalkonium chloride 10%	SU	10%	0.32	0.31	0.31	0.01	3.2
Cetylpyridinium bromide 6%	SU	100%	1.36	*	*		
Cetylpyridinium bromide 10%	SU	100%	*	*			
Polyethylene glycol 400	SU	100%	296.5	316.23	306.36	13.95	4.6
Mean							**23.3**
Median							**13.5**

* - Participating laboratory did not test the chemical because it determined that chemical was not compatible with the test system.

Table 7-17 Surfactant based formulations and mixtures - Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study.

Chemical	Formulation Tested	Conc. tested	MRD_{50} Values (mg/mL)		Mean MRD_{50} (mg/mL)	SD	CV (%)
			MA	CT AB			
Perfumed skin lotion	SU	100%	*	*			
Eye make-up remover	SU	100%	87.77	99.31	93.54	8.16	8.7
Hair dye base F#1	SU	100%	*	*			
Pump Deodorant	SU	5%	19.35	47.74	33.54	20.08	59.9
Emulsion antiperspirant	SU	100%	*	*			
Gel cleaner	SU	100%	5.68	5.47	5.58	0.15	2.6
Sunscreen SPF 15	SU	100%	*	*			
Hydrophilic ointment	SU	100%	*	*			
Hair conditioner	SU	100%	*	*			
Moisturiser with sunscreen	SU	100%	*	*			
Hair dye base form #3	SU	100%	*	*			
Polishing scrub	SU	100%	*	*			
Shampoo #1 normal	SU	100%	0.75	0.72	0.74	0.02	2.2
Hand cleaner	SU	100%	*	*			
Hand soap	SU	100%	*	*			
Shampoo - baby	SU	100%	2.51	2.15	2.33	0.25	10.8
Liquid soap #1	SU	100%	0.88	0.68	0.78	0.14	18.5
Shampoo antidandruff	SU	100%	*	*			
Shampoo 2-in-1	SU	100%	*	*			
Cleansing foam III	SU	100%	*	*			
Shower gel	SU	100%	*	*			
Skin cleaner	SU	100%	0.63	0.76	0.7	0.09	13.0
Mean							**16.5**
Median							**10.8**

* - Participating laboratory did not test the chemical because it determined that chemical was not compatible with the test system.

Table 7-18 Non-Surfactants, ingredients, and mixtures – Between-laboratories reproducibility of Cytosensor Microphysiometer results from COLIPA study.

Chemical	Formulation Tested	Conc. Tested	MRD_{50} Values (mg/mL)		Mean MRD_{50} (mg/mL)	SD	CV (%)
			MA	CT AB			
Blush		100%	*	*			
Eye liner		100%	*	*			
n-Butyl acetate		100%	*	*			
Imidazole		100%	18.84	26.03	22.43	5.09	22.7
Propylene glycol		100%	265.07	218.86	241.97	32.67	13.5
Glycerol	SO	100%	214.83	208.7	211.77	4.34	2.0
Ethyl acetate		100%	*	*			
Sodium hydroxide 1%	AL	1%	9.09	13.59	11.34	3.19	28.1
Isopropanol	SO	100%	52.59	124.51	88.55	50.86	57.4
Methyl ethyl ketone		1%	54.18	*	54.18		
Sunscreen lotion		10%	*	*			
Cologne		100%	*	*			
Eye shadow		100%	*	*			
Mascara		100%	*	*			
Hair styling lotion		100%	164.82	292.01	228.41	89.94	39.4
Mouthwash		100%	37.84	46.85	42.35	6.37	15.0
Toothpaste		100%	*	*			
Hair dye base form #2		100%	*	*			
Sodium hydroxide 10%	AL	6%	4.33	0.6	2.47	2.64	106.9
Trichloroacetic acid 30%	AC	30%	1.12	1.24	1.18	0.09	7.3
Mean							**32.5**
Median							**22.7**

* - Participating laboratory did not test the chemical because it determined that chemical was not compatible with the test system.

Table 7-19 Distribution of product categories for the interlaboratory reproducibility of the COLIPA study.

Product Categories	Number of products tested
Surfactants	22
Acids	1
Bases	2
Solvents	2
Other	15
Total	**41**

7.2 EpiOcular model

7.2.1 EpiOcular intralaboratory repeatability for antimicrobial cleaning products submitted by participating companies (within run and between experiments)

The within-run reproducibility can be assessed for some (15) of the formulations submitted for this study (Table 7-20). Studies were completed with at least 4 exposure times and on duplicate tissues for each exposure time. The distribution of product categories for the intralaboratory repeatability of the EpiOcular assay is shown in Table 7-21.

Table 7-20 EpiOcular intralaboratory repeatability both within run and between experiments

Code	Class	Assay Date	EPA Cat.	GHS Cat.	Conc.	ET_{50} (min)	Exp. Time (min)	Tissue 1 (% Survival)	Tissue 2 (% Survival)	Mean (%)	S.D.	CV (%)
H	AL	12/07/05	II	2A	Neat	9.4	20	28.4	26.7	**27.5**	1.2	4.5
							10	45.9	44.5	**45.2**	1.0	2.2
							5	82.9	86.9	**84.9**	2.9	3.4
							1	104.3	98.9	**101.6**	3.8	3.8
H	AL	04/05/06	II	2A	Neat	9.8	20	13.9	28.9	**21.4**	10.6	49.5
							12	37.4	22.2	**29.8**	10.8	36.0
							10	27.3	70.1	**48.7**	30.2	62.1
							5	76.8	80.9	**78.8**	2.8	3.6
							1	89.4	87.6	**88.5**	1.3	1.5
H	AL	04/05/06	II	2A	Neat	9.1	20	17.6	25.4	**21.5**	5.5	25.6
							12	15.1	29.9	**22.5**	10.5	46.7
							10	29.8	56.6	**43.2**	18.9	43.9
							5	78.3	81.9	**80.1**	2.5	3.1
							1	90.4	92.0	**91.2**	1.1	1.2
J	SU	12/07/05	III	2B	Neat	19.3	45	19.3	23.4	**21.4**	2.9	13.7
							20	46.1	48.4	**47.3**	1.6	3.4
							10	80.4	89.6	**85.0**	6.5	7.6
							5	95.9	93.8	**94.9**	1.5	1.5
K	RC	12/07/05	IV	NI	Neat	> 240	240	93.7	98.0	**95.9**	3.1	3.2
							90	99.6	109.6	**104.6**	7.0	6.7
							45	105.6	114.5	**110.1**	6.3	5.7
							20	98.8	107.9	**103.4**	6.4	6.2
P	AL	12/07/05	IV	NI	Neat	125.8	240	21.0	27.4	**24.2**	4.5	18.7
							90	57.6	58.5	**58.1**	0.7	1.1
							45	88.4	92.1	**90.3**	2.6	2.9
							20	126.8	127.9	**127.4**	0.8	0.6
P	AL	04/05/06	IV	NI	Neat	74.0	240	9.2	16.2	**12.7**	5.0	39.2
							90	36.9	37.6	**37.3**	0.5	1.4
							45	72.6	73.4	**73.0**	0.5	0.7
							20	125.1	119.8	**122.5**	3.7	3.1

4811

Code	Class	Assay Date	EPA Cat.	GHS Cat.	Conc.	ET_{50} (min)	Exp. Time (min)	Tissue 1 (% Survival)	Tissue 2 (% Survival)	Mean (%)	S.D.	CV (%)
R	SU	12/07/05	IV	NI	Neat	> 240	240	83.5	82.0	**82.8**	1.1	1.3
							90	96.8	80.8	**88.8**	11.3	12.8
							45	110.9	97.2	**104.1**	9.7	9.3
							20	92.8	99.1	**96.0**	4.4	4.6
T	AC	12/07/05	IV	NI	Neat	31.6	240	8.3	8.8	**8.5**	0.4	4.1
							90	8.5	9.5	**9.0**	0.7	7.3
							45	19.6	20.7	**20.1**	0.8	4.0
							20	75.7	76.2	**76.0**	0.3	0.4
W	SU	12/07/05	IV	NI	Neat	39.6	45	41.1	48.8	**44.9**	5.4	12.1
							20	62.6	74.3	**68.5**	8.3	12.1
							10	84.4	90.8	**87.6**	4.5	5.2
							5	96.2	104.8	**100.5**	6.1	6.1
W	SU	04/05/06	IV	NI	Neat	43.3	90	14.9	29.1	**22.0**	10.0	45.4
							45	52.2	43.9	**48.0**	5.9	12.2
							20	78.5	77.6	**78.1**	0.6	0.8
							5	91.0	88.7	**89.8**	1.6	1.8
V	SU	12/07/05	IV	NL	Neat	12.0	20	25.8	25.0	**25.4**	0.6	2.3
							10	53.1	58.8	**56.0**	4.0	7.2
							5	68.4	69.1	**68.8**	0.5	0.7
							1	103.6	97.1	**100.3**	4.6	4.6
AT	RC	12/07/05	I	1	Neat	<1	20	0.8	0.8	**0.8**	0.0	5.9
							10	0.7	0.6	**0.6**	0.1	14.9
							5	4.9	6.9	**5.9**	1.4	23.6
							1	9.0	14.3	**11.6**	3.7	32.2
AU	RC	12/07/05	I	1	Neat	<1	20	0.9	0.9	**0.9**	0.0	0.0
							10	0.8	1.0	**0.9**	0.1	10.5
							5	3.8	2.7	**3.2**	0.8	23.8
							1	5.1	7.7	**6.4**	1.9	29.3
AV	RC	12/07/05	I	1	Neat	<1	20	1.1	0.8	**0.9**	0.1	15.2
							10	1.1	1.8	**1.5**	0.5	32.9
							5	2.4	1.9	**2.1**	0.4	18.0
							1	3.8	8.1	**6.0**	3.0	50.6
BB	SO	12/07/05	IV	SCNM	Neat	>240	240	96.7	101.6	**99.2**	3.5	3.5
							90	104.6	106.6	**105.6**	1.4	1.3
							45	90.7	101.6	**96.1**	7.7	8.0
							20	104.0	96.4	**100.2**	5.4	5.4
BK		12/07/05	III	NI	Neat	9.4	20	37.8	32.8	**35.3**	3.5	10.1
							10	40.5	49.1	**44.8**	6.1	13.6
							5	82.7	91.7	**87.2**	6.4	7.3
							1	107.5	106.2	**106.8**	0.9	0.9
BM	SO	12/07/05	IV	NI	Neat	4.9	20	14.5	16.3	**15.4**	1.2	8.0
							10	18.4	18.4	**18.4**	0.0	0.0
							5	43.0	53.7	**48.4**	7.6	15.7
							1	97.6	99.7	**98.7**	1.4	1.4

Code	Class	Assay Date	EPA Cat.	GHS Cat.	Conc.	ET_{50} (min)	Exp. Time (min)	Tissue 1 (% Survival)	Tissue 2 (% Survival)	Mean (%)	S.D.	CV (%)
BL	SO	12/07/05	IV	NI	Neat	6.7	20	37.3	34.7	**36.0**	1.8	5.1
							10	34.7	29.4	**32.1**	3.8	11.8
							5	60.7	58.0	**59.4**	1.9	3.2
							1	100.8	104.5	**102.6**	2.6	2.6

Table 7-21 Distribution of product categories for the intralaboratory repeatability of the EpiOcular assay.

Product Categories	Number of products tested
Surfactants	4
Acids	1
Bases	2
Solvents	3
Oxidizers	4
Other	1
Total	**15**

7.2.2 EpiOcular intralaboratory reproducibility (between experiments)

Intralaboratory reproducibility for EpiOcular can be estimated from the repeat testing of a single material (0.3% Triton X-100) over a nine year period (Table 7-22). Percent CV for the two laboratories combined was 20.7% and for a single laboratory (IIVS) was 22.2% (eight years only). The standard deviation range for the 0.3% Triton X-100 over a nine year period is described in Table 7-23.

Table 7-22 Intralaboratory reproducibility of EpiOcular tissue over a nine year period from 1997 through 2005 for two different laboratories.

Test Material	Mean ET_{50} Value (min)	Standard Deviation (min)	CV (%)
0.3% Triton X-100 (Combined data from MatTek and IIVS)	26.1	5.4	20.7
0.3% Triton X-100 (IIVS only-through Oct., 2004)	27.0	6.0	22.2

Table 7-23 Standard deviation range for 0.3% Triton X-100 for EpiOcular tissue over a nine year period.

SD Range	1997 (%)	1998 (%)	1999 (%)	2000 (%)	2001 (%)	2002 (%)	2003 (%)	2004 (%)	2005 (%)	1997-2005 YTD (%)
0.0 to 0.5	41	52	36	29	35	32	36	33	47	38
0.5 to 1.0	35	26	31	25	36	22	26	27	35	29
1.0 to 1.5	20	17	24	27	20	31	25	19	15	22
0.0 to 1.5	95	95	92	81	91	85	87	79	97	89
1.5 to 2.0	5	5	8	18	9	15	13	21	3	11
# Production Lots	>50	>50	>50	>50	>50	>50	>50	>50	>50	>500
Average ET_{50} (min)	22.9	25.0	22.1	20.7	22.9	22.5	24.1	22.2	24.77	23.00

Additional information on the intralaboratory reproducibility of EpiOcular can be found in the BRD prepared for ECVAM.

7.2.3 EpiOcular interlaboratory reproducibility

Data on interlaboratory reproducibility can be found in the BRD prepared for ECVAM. Two specific examples from two phases of a validation study conducted for Colgate-Palmolive are shown below. This validation study was conducted with surfactants and surfactant-containing products to investigate a different prediction model than is presented in this BRD. The prediction model tested in Phase II (Table 7-24) and Phase III (Table 7-25) is based on Draize MAS scores, and consequently the reproducibility comparisons are based on predicted MAS scores, not directly on ET_{50} values. However, the values given do reflect the reproducibility that can be expected using ET_{50} values as is done in this BRD. It can be seen that the mean %CV in Phase II (4 laboratories) was 18.1% and in Phase III (2 labs) was 11.8%. The distribution of product categories for the interlaboratory reproducibility of the EpiOcular assay is shown in Table 7-26.

Table 7-24 Interlaboratory reproducibility of four laboratories in the Colgate-Palmolive Phase II validation study.

Test Material	Formulation Type	Predicted Draize Score				Average	SD	CV (%)
		Lab 1	Lab 2	Lab 3	Lab 4			
Shampoo #1 (2 in 1)	SU	14.5	16.7	18.9	18.2	17.1	1.9	11.4
Shampoo #2 (Regular)	SU	28.6	27.2	31.3	30.6	29.4	1.9	6.4
Shampoo #3 (Regular)	SU	29.7	25.5	30.5	32.4	29.5	2.9	9.9
Dishwashing Liquid	SU	79.7	58.2	35.1	97	67.5	26.8	39.7
All purpose cleaner	SU	23.5	20.5	21.2	28.1	23.3	3.4	14.7
Disinfectant cleaner	SU	40	36.4	45.7	34.8	39.2	4.8	12.3
Sodium linear alkylbenzene sulfonate	SU	36.1	39.8	38.4	40.5	38.7	1.9	5.0
30% Dimethyltetradecylamine oxide	SU	58.3	88.3	84.7	50.2	70.4	19.0	27.0
1.5% branched alkyldimethylamine	SU	21.6	22.3	26.4	23.8	23.5	2.1	9.0
PPG-5 Ceteth-20	SU	3.1	3.4	4.8	3.6	3.7	0.7	20.0
C9-11 Alcohol ethoxylate EO6:1	SU	61.7	53.7	33.7	32.1	45.3	14.7	32.5
C12-14 Alcohol ethoxylate 2EO	SU	6	4.9	9.9	7.1	7.0	2.1	30.8
C12-16 Alcohol ethoxylate 3EO	SU	8.7	10.3	11.2	9.8	10.0	1.0	10.4
2.46% Lauryl hydroxysultaine	SU	24.2	25.1	27.3	23.5	25.0	1.7	6.6
10% Polyoxyethylene (10) oleyl ether	SU	1.8	3.1	3.1	4.3	3.1	1.0	33.2
3.2% Benzalkonium chloride	SU	71.8	60.3	78.8	62.2	68.3	8.6	12.6
36% Sodium methyl 2-sulfonate & disodium 2-sulfolaurate	SU	35.4	39.8	36.2	34.8	36.6	2.2	6.1
2.4% Imidazolium compound	SU	1.8	2.5	2.8	1.8	2.2	0.5	22.7
C12-15 Alcohol ethoxylate EO7:1	SU	6.2	5	8.7	10.7	7.7	2.6	33.4
Mean CV (%)								**18.1**
Median CV (%)								**12.6**

Table 7-25 Interlaboratory reproducibility of two laboratories in the Colgate-Palmolive Phase III validation study.

Test Material	Formulation Type	Concentration Tested	Predicted Draize Score		Average	SD	CV (%)
			Lab 1	Lab 2			
1-decanaminium, N-decyl-N,N-dimethyl, Cl-	SU	50%	97	97	97.0	0.0	0.0
20% Myristalkonium chloride/ 20% Quaternium-14	SU	100%	97	92.2	94.6	3.4	3.6
Alkyldimethyl benzyl ammonium Cl-	SU	5%	60	84	72.0	17.0	23.6
Ammonium lauryl sulfate	SU	12%	25.5	25.4	25.5	0.1	0.3
Ammonium lauryl sulfate	SU	28%	34.3	29.4	31.9	3.5	10.9
Ammonium nonoxyl-4 sulfate	SU	10%	17.5	19.4	18.5	1.3	7.3
Behentrimonium methosulfate & cetearyl alcohol	SU	100%	1.8	1.8	1.8	0.0	0.0

4852

Test Material	Formulation Type	Concentration Tested	Predicted Draize Score		Average	SD	CV (%)
			Lab 1	Lab 2			
Benzalkonium chloride	SU	0.10%	1.8	1.8	1.8	0.0	0.0
Benzalkonium chloride	SU	0.50%	18.1	24.2	21.2	4.3	20.4
Benzalkonium chloride	SU	1%	27.9	35.5	31.7	5.4	17.0
Benzalkonium chloride	SU	2.50%	66.4	65.8	66.1	0.4	0.6
Benzalkonium chloride	SU	5%	68.3	96.5	82.4	19.9	24.2
Benzalkonium chloride	SU	10%	90.2	97	93.6	4.8	5.1
Benzethonium chloride	SU	3.20%	42.1	56.5	49.3	10.2	20.7
Benzethonium chloride	SU	1.00%	29.2	41.7	35.5	8.8	24.9
Branched alkyldimethylamine	SU	1.50%	16.8	20.4	18.6	2.5	13.7
Branched alkyldimethylamine	SU	30%	97	97	97.0	0.0	0.0
C10-12 Alcohol ethoxylate (PO)	SU	100%	87.6	80.6	84.1	4.9	5.9
Ceteareth-12	SU	100%	1.8	4.1	3.0	1.6	55.1
Cetrimonium chloride	SU	2.50%	22.2	19.7	21.0	1.8	8.4
Cetyl alcohol	SU	100%	1.8	1.8	1.8	0.0	0.0
Cetylpyridinium bromide	SU	10%	25.2	24.8	25.0	0.3	1.1
Cetylpyridinium bromide	SU	0.1%	1.8	1.8	1.8	0.0	0.0
Cetylpyridinium bromide	SU	1%	10.2	17.8	14.0	5.4	38.4
Cocamidopropyl betaine	SU	10%	23.7	32.6	28.2	6.3	22.4
Cocamidopropyl betaine	SU	30%	44.1	46.4	45.3	1.6	3.6
Decyl glucoside	SU	10%	21.1	23	22.1	1.3	6.1
Didecyldimethyl ammonium chloride (DDAC)	SU	1%	32.5	39.9	36.2	5.2	14.5
Didecyldimethyl ammonium chloride (DDAC)	SU	3.20%	62.9	72.9	67.9	7.1	10.4
Didecyldimethyl ammonium chloride (DDAC)	SU	5%	9	14.8	11.9	4.1	34.5
Lauryl glucoside	SU	12%	2.5	3.4	3.0	0.6	21.6
Myristalkonium chloride/Quaternium-14/Ethanol	SU	3%	40.2	59.5	49.9	13.6	27.4
Myristalkonium chloride/Quaternium-14/Ethanol	SU	20%	62.8	97	79.9	24.2	30.3
PPG-5-Ceteth 20	SU	100%	1.8	3.5	2.7	1.2	45.4
Quaternium-18	SU	100%	1.8	1.8	1.8	0.0	0.0
Shampoo #4	SU	10%	14.3	15.3	14.8	0.7	4.8
Sodium C14-16 olefin sulfonate	SU	10%	19.2	20	19.6	0.6	2.9
Sodium ether sulfate 3EO	SU	30%	30.6	30.5	30.6	0.1	0.2
Sodium laureth sulfate	SU	12%	18.5	21	19.8	1.8	9.0
Sodium laureth sulfate	SU	25%	23.7	27.1	25.4	2.4	9.5
Sodium lauroyl sarcosinate	SU	10%	24.8	23.2	24.0	1.1	4.7
Sodium lauroyl sarcosinate	SU	30%	33	35.1	34.1	1.5	4.4
Sodium lauryl sulfate	SU	3%	23.2	24	23.6	0.6	2.4
Sodium lauryl sulfate	SU	10%	30.3	33.4	31.9	2.2	6.9
Sodium lauryl sulfate	SU	15%	34.7	36.8	35.8	1.5	4.2
Sodium lauryl sulfate	SU	20%	39.6	41.8	40.7	1.6	3.8
Sodium lauryl sulfate	SU	30%	39.6	47.3	43.5	5.4	12.5
Sodium methyl 2-sulfonate & disodium 2-sulfolaurate	SU	39%	33.4	35.3	34.4	1.3	3.9

Test Material	Formulation Type	Concentration Tested	Predicted Draize Score		Average	SD	CV (%)
			Lab 1	Lab 2			
TEA-lauryl sulfate	SU	20%	26.5	32.1	29.3	4.0	13.5
Triton X-100	SU	1%	9.7	12.1	10.9	1.7	15.6
Triton X-100	SU	2.50%	24.1	22.8	23.5	0.9	3.9
Triton X-100	SU	5%	36.6	46	41.3	6.6	16.1
Triton X-100	SU	10%	51.8	53.7	52.8	1.3	2.5
Triton X-100	SU	20%	50.2	63.8	57.0	9.6	16.9
Mean CV (%)							**11.8**
Median CV (%)							**7.1**

Table 7-26 Distribution of product categories for the interlaboratory reproducibility of the EpiOcular assay.

Product Categories	Number of products tested
Surfactants	73 (including unique dilutions of products)
Acids	0
Bases	0
Solvents	0
Other	0
Total	**73**

7.3 BCOP assay

7.3.1 BCOP intralaboratory repeatability

7.3.1.1 BCOP within-run reproducibility for antimicrobial cleaning products data

Data from 75 runs (255 corneas; 3-5 corneas per run) of the BCOP assay for antimicrobial cleaning products are presented in Table 7-27. This table displays, for each test material run, the mean value and the mean %CV for the three main parameters of the BCOP assay – opacity, permeability, and *in vitro* score. Also presented is the overall mean of the %CV for each of the parameters. The distribution of product categories for the within-run reproducibility of the BCOP assay is shown in Table 7-28.

BCOP scores in the very low range (arbitrarily set in this BRD as ≤10) can often generate high %CV's, but these %CV's have no practical use in evaluating the utility of the assay since the individual measurements may only differ by one or two units on a scale that reaches into the hundreds. For example, the three opacity readings of 0, 2, 3 yield an extraordinarily high %CV of 91.7% even though the three values are essentially the same when the overall scale is considered. Therefore, %CV's from materials whose average *in vitro* score is ≤10 (first section of Table 7-27) will be considered separately from the %CV's of those materials whose average *in vitro* score is >10 (second section of Table 7-27)..

Table 7-27 indicates that there is a large difference between the %CV's for BCOP values where the average *In Vitro* Score is ≤ 10 and those cases where it is >10. The average %CV's for opacity values, permeability values and *in vitro* scores, in the first case, are 266%, 167.1% and 66.4%, respectively. However, in the second case they are much lower: 27.9%, 24.1% and 18.3%, respectively. It is clear that where small opacity values are recorded (the first case), the percent CV is really meaningless as a way of judging reproducibility. However, in the second case with higher numbers the average %CV's indicate that the BCOP assay has a high within run reproducibility.

Table 7-27 BCOP within run reproducibility

Materials with an average *In Vitro* Score ≤10										
Test Material Code	**Formulation Type**	**Cornea Number**	**Opacity**	**Permeability**	***In vitro* Score**	***In vivo* GHS Cat**	***In vivo* EPA Cat**	**Opacity CV**	**Perm. CV**	***In vitro* Score CV**
H	AL	43	-0.3	0.348	4.9	Non-irritant	Category III	1732.1%	61.8%	67.8%
		44	-0.3	0.444	6.4					
		45	0.7	1.047	16.4					
		avg	0.0	0.613	9.2					
		S.D.	0.6	0.379	6.3					
		% CV	1732.1%	61.8%	67.8%					
I	SU	34	-1.3	0.088	0.0	Non-irritant	Category III	57.7%	13.4%	127.7%
		35	-1.3	0.111	0.3					
		36	-0.3	0.113	1.4					
		avg	-1.0	0.104	0.6					
		S.D.	0.6	0.014	0.7					
		% CV	57.7%	13.4%	127.7%					
J	SU	23	6.3	0.085	7.6	Non-irritant	Category III	32.7%	24.1%	22.5%
		24	3.3	0.098	4.8					
		25	4.3	0.134	6.3					
		avg	4.7	0.105	6.2					
		S.D.	1.5	0.025	1.4					
		% CV	32.7%	24.1%	22.5%					
K	RC	21	0.3	-0.001	0.3	Non-irritant	Category IV	0.0%	692.8%	10.2%
		22	0.3	0.003	0.4					
		24	0.3	-0.001	0.3					
		avg	0.3	0.000	0.3					
		S.D.	0.0	0.002	0.0					
		% CV	0.0%	692.8%	10.2%					
L	SU	17	3.7	0.035	4.2	Non-irritant	Category III	21.4%	76.6%	28.6%
		18	5.7	0.106	7.3					
		19	4.7	0.028	5.1					
		avg	4.7	0.056	5.5					
		S.D.	1.0	0.043	1.6					
		% CV	21.4%	76.6%	28.6%					

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
O	SU	14	7.3	0.030	7.8	Non-irritant	Category IV	17.3%	26.9%	14.0%
		15	7.3	0.032	7.8					
		16	5.3	0.048	6.1					
		avg	6.7	0.037	7.2					
		S.D.	1.2	0.010	1.0					
		% CV	17.3%	26.9%	14.0%					
P	AL	29	1.3	0.019	1.6	Non-irritant	Category IV	57.7%	168.8%	61.9%
		30	1.3	0.001	1.3					
		31	0.3	0.000	0.3					
		avg	1.0	0.006	1.1					
		S.D.	0.6	0.011	0.7					
		% CV	57.7%	168.8%	61.9%					
R	SU	35	0.3	0.019	0.6	Non-irritant	Category IV	1732.1%	60.5%	261.8%
		36	0.3	0.005	0.4					
		46	-0.7	0.022	-0.4					
		avg	0.0	0.015	0.2					
		S.D.	0.6	0.009	0.5					
		% CV	1732.1%	60.5%	261.8%					
T	AC	27	-0.3	0.008	-0.2	Non-irritant	Category IV	103.9%	21.5%	96.8%
		28	2.7	0.012	2.8					
		29	2.7	0.009	2.8					
		avg	1.7	0.010	1.8					
		S.D.	1.7	0.002	1.8					
		% CV	103.9%	21.5%	96.8%					
U	SU	26	2.7	0.050	3.4	Non-irritant	Category IV	24.7%	55.1%	29.9%
		27	2.7	0.120	4.5					
		28	1.7	0.050	2.4					
		avg	2.3	0.073	3.4					
		S.D.	0.6	0.040	1.0					
		% CV	24.7%	55.1%	29.9%					

4892

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
W	SU	45	3.7	0.074	4.8	Non-irritant	Category IV	37.1%	5.3%	30.6%
		47	6.7	0.073	7.8					
		48	3.7	0.067	4.7					
		avg	4.7	0.071	5.7					
		S.D.	1.7	0.004	1.8					
		% CV	37.1%	5.3%	30.6%					
AF	AC	34	13.0	0.013	13.2	Non-irritant	Category IV	34.4%	33.3%	32.4%
		37	8.0	0.021	8.3					
		38	7.0	0.026	7.4					
		avg	9.3	0.020	9.6					
		S.D.	3.2	0.007	3.1					
		% CV	34.4%	33.3%	32.4%					
BB	SO	25	1.0	-0.004	0.9	SCNM	Category IV	132.3%	-1249.0%	135.5%
		26	0.0	-0.002	0.0					
		28	5.0	0.004	5.1					
		avg	2.0	0.000	2.0					
		S.D.	2.6	0.004	2.7					
		% CV	132.3%	-1249.0%	135.5%					
BK	SO	29	2.0	0.414	8.2	Non-irritant	Category III	88.2%	60.2%	45.9%
		30	6.0	0.178	8.7					
		31	1.0	0.143	3.1					
		avg	3.0	0.245	6.7					
		S.D.	2.6	0.147	3.1					
		% CV	88.2%	60.2%	45.9%					
BL	SO	14	1.0	0.294	5.4	Non-irritant	Category IV	142.0%	32.8%	51.8%
		16	0.0	0.211	3.2					
		17	7.0	0.151	9.3					
		avg	2.7	0.219	6.0					
		S.D.	3.8	0.072	3.1					
		% CV	142.0%	32.8%	51.8%					

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
CG	AL	12	5.3	0.027	5.7	Non-irritant	Category IV	41.7%	90.7%	44.3%
		13	2.3	0.001	2.4					
		14	3.3	0.014	3.5					
		avg	3.7	0.014	3.9					
		S.D.	1.5	0.013	1.7					
		% CV	41.7%	90.7%	44.3%					
Mean CV for materials with an i*n vitro* score less than 10								**266%**	**167.1%**	**66.4%**
Materials with an average *In Vitro* Score > 10										
Test Material Code	**Formulation Type**	**Cornea Number**	**Opacity**	**Permeability**	***In vitro* Score**	***In vivo* GHS Cat**	***In vivo* EPA Cat**	**Opacity CV**	**Perm. CV**	***In vitro* Score CV**
H	AL	15	0.3	0.614	9.5	Non-irritant	Category III	57.7%	27.4%	27.3%
		17	1.3	0.914	15.0					
		18	1.3	0.560	9.7					
		avg	1.0	0.696	11.4					
		S.D.	0.6	0.191	3.1					
		% CV	57.7%	27.4%	27.3%					
H	AL	48	2.7	0.401	8.7	Non-irritant	Category III	66.6%	41.6%	46.1%
		49	2.7	0.783	14.4					
		50	7.7	0.999	22.7					
		avg	4.3	0.728	15.2					
		S.D.	2.9	0.303	7.0					
		% CV	66.6%	41.6%	46.1%					
H	AL	37	2.7	0.698	13.1	Non-irritant	Category III	60.0%	16.6%	9.4%
		38	1.7	0.673	11.8					
		40	0.7	0.902	14.2					
		avg	1.7	0.758	13.0					
		S.D.	1.0	0.126	1.2					
		% CV	60.0%	16.6%	9.4%					

4893

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
H	AL	50	1.3	0.911	15.0	Non-irritant	Category III	43.3%	37.2%	26.7%
		51	3.3	0.425	9.7					
		52	3.3	0.899	16.8					
		avg	2.7	0.745	13.8					
		S.D.	1.2	0.277	3.7					
		% CV	43.3%	37.2%	26.7%					
H	AL	29	2.7	1.101	19.2	Non-irritant	Category III	86.6%	15.4%	31.3%
		32	2.7	0.884	15.9					
		33	10.7	1.205	28.7					
		avg	5.3	1.064	21.3					
		S.D.	4.6	0.164	6.7					
		% CV	86.6%	15.4%	31.3%					
Q	SU	42	8.7	0.223	12.0	Non-irritant	Category IV	6.4%	66.4%	26.3%
		43	8.7	0.150	10.9					
		44	9.7	0.525	17.5					
		avg	9.0	0.299	13.5					
		S.D.	0.6	0.199	3.6					
		% CV	6.4%	66.4%	26.3%					
V	SU	19	4.7	1.090	21.0	Non-irritant	Category IV	11.5%	10.3%	5.1%
		20	5.7	0.928	19.6					
		21	4.7	1.132	21.6					
		avg	5.0	1.050	20.8					
		S.D.	0.6	0.108	1.1					
		% CV	11.5%	10.3%	5.1%					
X	RC	19	7.7	4.772	79.3	Category 2A	Category I	39.7%	4.4%	5.2%
		21	8.7	5.207	86.8					
		22	3.7	5.067	79.7					
		avg	6.7	5.016	81.9					
		S.D.	2.6	0.222	4.2					
		% CV	39.7%	4.4%	5.2%					

4894

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
Z	SO	39	6.0	1.358	26.4	Category 2A	Category II	24.7%	17.8%	14.7%
		41	6.0	1.942	35.1					
		43	9.0	1.627	33.4					
		avg	7.0	1.642	31.6					
		S.D.	1.7	0.292	4.6					
		% CV	24.7%	17.8%	14.7%					
AQ	RC	11	4.7	5.055	80.5	Category 1	Category I	57.7%	12.1%	8.1%
		12	4.7	5.875	92.8					
		13	11.7	4.640	81.3					
		avg	7.0	5.190	84.9					
		S.D.	4.0	0.628	6.9					
		% CV	57.7%	12.1%	8.1%					
AS	RC	27	3.7	4.860	76.6	Category 1	Category I	61.1%	18.7%	14.0%
		28	3.7	5.905	92.2					
		29	9.7	4.065	70.6					
		avg	5.7	4.944	79.8					
		S.D.	3.5	0.923	11.2					
		% CV	61.1%	18.7%	14.0%					
AT	RC	34	2.7	5.870	90.7	Category 1	Category I	19.2%	9.9%	8.8%
		35	2.7	5.760	89.1					
		36	3.7	4.880	76.9					
		avg	3.0	5.504	85.6					
		S.D.	0.6	0.543	7.6					
		% CV	19.2%	9.9%	8.8%					
AW	RC	29	4.7	4.000	64.7	Category 1	Category I	13.3%	26.1%	25.1%
		30	3.7	3.775	60.3					
		34	4.7	5.950	93.9					
		avg	4.3	4.575	73.0					
		S.D.	0.6	1.196	18.3					
		% CV	13.3%	26.1%	25.1%					

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
BD	SO	16	6.7	1.210	24.9	Non-irritant	Category III	170.8%	12.0%	31.9%
		17	-1.3	0.994	13.6					
		18	1.7	0.984	16.5					
		avg	2.4	1.063	18.3					
		S.D.	4.0	0.128	5.8					
		% CV	170.8%	12.0%	31.9%					
BP	SO	11	7.0	0.848	19.7	Non-irritant	Category IV	7.5%	20.7%	11.3%
		12	8.0	0.865	21.0					
		13	8.0	0.583	16.7					
		avg	7.7	0.765	19.1					
		S.D.	0.6	0.158	2.2					
		% CV	7.5%	20.7%	11.3%					
A	SU	25	175.7	2.925	219.5	Category 1	Category I	6.0%	15.3%	5.3%
		26	166.7	2.245	200.3					
		28	155.7	3.005	200.7					
		avg	166.0	2.725	206.9					
		S.D.	10.0	0.418	11.0					
		% CV	6.0%	15.3%	5.3%					
B	SU	11	138.7	0.946	152.9	Category 1	Category I	2.2%	7.4%	1.5%
		12	135.7	0.932	149.6					
		13	141.7	0.824	154.0					
		avg	138.7	0.901	152.2					
		S.D.	3.0	0.067	2.3					
		% CV	2.2%	7.4%	1.5%					
C	RC	47	16.3	0.460	23.2	Category 1	Category I	21.7%	47.0%	30.9%
		48	16.3	0.624	25.7					
		49	23.3	1.124	40.2					
		avg	18.7	0.736	29.7					
		S.D.	4.0	0.346	9.2					
		% CV	21.7%	47.0%	30.9%					

4895

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
D	AC	15	179.7	0.001	179.7	Category 1	Category I	4.3%	43.3%	4.3%
		16	187.7	0.003	187.7					
		18	195.7	0.003	195.7					
		avg	187.7	0.003	187.7					
		S.D.	8.0	0.001	8.0					
		% CV	4.3%	43.3%	4.3%					
E	SU	15	172.7	2.196	205.6	Category 1	Category I	1.8%	21.2%	4.2%
		16	168.7	1.442	190.3					
		19	166.7	1.741	192.8					
		avg	169.3	1.793	196.2					
		S.D.	3.1	0.380	8.2					
		% CV	1.8%	21.2%	4.2%					
F	RC	32^	22.7	2.053	53.5	Category 1	Category I	81.0%	167.7%	73.8%
		35	513.7	0.044	514.3					
		37	513.7	0.001	513.7					
		avg	350.0	0.699	360.5					
		S.D.	283.5	1.173	265.9					
		% CV	81.0%	167.7%	73.8%					
F	RC	22	410.7	0.157	413.0	Category 1	Category I	145.9%	78.5%	118.5%
		24	25.7	1.848	53.4					
		26	22.7	2.273	56.8					
		avg	153.0	1.426	174.4					
		S.D.	223.2	1.119	206.7					
		% CV	145.9%	78.5%	118.5%					
G	SU	12	88.7	3.530	141.6	Category 1	Category I	12.8%	4.0%	9.5%
		13	96.7	3.680	151.9					
		14	74.7	3.395	125.6					
		avg	86.7	3.535	139.7					
		S.D.	11.1	0.143	13.2					
		% CV	12.8%	4.0%	9.5%					

4896

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
M	SU	32	25.3	1.334	45.3	Non-irritant	Category III	20.3%	30.1%	17.2%
		35	38.3	1.281	57.5					
		37	32.3	2.131	64.3					
		avg	32.0	1.582	55.7					
		S.D.	6.5	0.476	9.6					
		% CV	20.3%	30.1%	17.2%					
N	RC	15	6.7	0.997	21.6	Non-irritant	Category III	164.5%	84.4%	148.2%
		16	7.7	0.991	22.5					
		17	413.7	0.017	413.9					
		avg	142.7	0.668	152.7					
		S.D.	234.7	0.564	226.2					
		% CV	164.5%	84.4%	148.2%					
S	AC	22	12.7	0.302	17.2	Non-irritant	Category IV	11.8%	18.7%	10.3%
		23	14.7	0.417	20.9					
		24	11.7	0.434	18.2					
		avg	13.0	0.385	18.8					
		S.D.	1.5	0.072	1.9					
		% CV	11.8%	18.7%	10.3%					
Y	RC	28	35.7	2.612	74.9	Category 2A	Category II	9.4%	20.6%	9.7%
		29	29.7	3.497	82.1					
		33	31.7	2.397	67.6					
		avg	32.3	2.836	74.9					
		S.D.	3.1	0.583	7.3					
		% CV	9.4%	20.6%	9.7%					
AB	SU	27	85.0	1.434	106.5	Category 1	Category I	24.0%	13.5%	16.0%
		28	54.0	1.770	80.5					
		29	62.0	1.394	82.9					
		avg	67.0	1.532	90.0					
		S.D.	16.1	0.207	14.4					
		% CV	24.0%	13.5%	16.0%					

4897

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AC	AC	25	113.0	1.289	132.3	Category 1	Category I	32.2%	31.7%	32.0%
		26	150.0	1.941	179.1					
		27	77.0	1.067	93.0					
		avg	113.3	1.432	134.8					
		S.D.	36.5	0.454	43.1					
		% CV	32.2%	31.7%	32.0%					
AD	SU	18	89.0	1.409	110.1	Category 1	Category I	3.1%	2.9%	2.3%
		19	94.0	1.338	114.1					
		20	94.0	1.406	115.1					
		avg	92.3	1.384	113.1					
		S.D.	2.9	0.040	2.6					
		% CV	3.1%	2.9%	2.3%					
AE	AL	17	43.0	1.455	64.8	Category 1	Category I	3.7%	23.8%	9.7%
		18	42.0	2.126	73.9					
		20	40.0	1.424	61.4					
		avg	41.7	1.668	66.7					
		S.D.	1.5	0.397	6.5					
		% CV	3.7%	23.8%	9.7%					
AG	AL	6	340.7	3.487	393.0	Category 1	Category I	3.6%	6.7%	2.7%
		7	343.0	3.217	391.3					
		8	329.3	3.192	377.2					
		9	363.7	2.887	407.0					
		10	344.3	3.127	391.2					
		avg	344.2	3.182	391.9					
		S.D.	12.4	0.215	10.6					
		% CV	3.6%	6.7%	2.7%					

4898

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AH	AL	6	278.3	1.869	306.3	Category 1	Category I	29.1%	8.6%	25.7%
		7	111.3	1.994	141.2					
		8	254.0	2.189	286.8					
		9	243.3	2.214	276.5					
		10	240.0	1.844	267.7					
		avg	225.4	2.022	255.7					
		S.D.	65.5	0.174	65.6					
		% CV	29.1%	8.6%	25.7%					
AI	AL	6	325.3	2.498	362.8	Category 1	Category I	4.3%	16.0%	5.2%
		7	299.0	2.478	336.2					
		8	306.3	2.108	337.9					
		9	332.0	3.238	380.6					
		10	314.3	2.773	355.9					
		avg	315.4	2.619	354.7					
		S.D.	13.5	0.419	18.4					
		% CV	4.3%	16.0%	5.2%					
AJ	AL	1	289.7	2.289	324.0	Category 1	Category I	9.8%	14.0%	9.4%
		2	312.7	2.234	346.2					
		3	354.3	1.944	383.5					
		4	300.7	1.999	330.7					
		5	360.0	2.734	401.0					
		avg	323.5	2.240	357.1					
		S.D.	31.9	0.313	33.7					
		% CV	9.8%	14.0%	9.4%					

4899

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AK	AL	16	401.7	2.869	444.7	Category 1	Category I	2.0%	12.6%	1.9%
		17	396.0	2.404	432.1					
		18	410.7	2.034	441.2					
		19	417.3	2.439	453.9					
		20	409.7	2.659	449.6					
		avg	407.1	2.481	444.3					
		S.D.	8.3	0.312	8.4					
		% CV	2.0%	12.6%	1.9%					
AL	AL	6	325.3	1.818	352.6	Category 2A	Category I	16.7%	20.4%	13.9%
		7	325.7	2.918	369.5					
		8	222.7	3.163	270.1					
		9	349.7	2.853	392.5					
		21	348.3	2.363	383.7					
		avg	314.3	2.623	353.7					
		S.D.	52.6	0.536	49.1					
		% CV	16.7%	20.4%	13.9%					
AM	SO	1	89.0	2.267	123.0	Category 1	Category I	6.6%	19.5%	9.4%
		2	103.0	2.517	140.8					
		3	98.0	2.612	137.2					
		4	96.0	1.887	124.3					
		5	105.7	3.212	153.9					
		avg	98.3	2.499	135.8					
		S.D.	6.5	0.487	12.7					
		% CV	6.6%	19.5%	9.4%					

4900

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AN	AL	1	88.0	1.173	105.6	Category 1	Category I	12.0%	16.0%	11.7%
		2	98.3	1.418	119.6					
		3	105.0	1.058	120.9					
		4	79.7	0.973	94.3					
		5	106.7	1.363	127.1					
		avg	95.5	1.197	113.5					
		S.D.	11.5	0.191	13.3					
		% CV	12.0%	16.0%	11.7%					
AO	AL	11	176.0	2.594	214.9	Category 1	Category I	7.9%	6.1%	6.3%
		12	159.7	2.314	194.4					
		13	192.7	2.289	227.0					
		14	194.7	2.244	228.4					
		15	179.3	2.459	216.2					
		avg	180.5	2.380	216.2					
		S.D.	14.2	0.144	13.6					
		% CV	7.9%	6.1%	6.3%					
AP	AL	16	358.3	3.431	409.8	Category 1	Category I	4.2%	14.2%	4.3%
		17	360.3	2.421	396.6					
		18	343.0	2.966	387.5					
		19	325.0	2.826	367.4					
		20	353.7	3.421	405.0					
		avg	348.1	3.013	393.3					
		S.D.	14.5	0.427	16.8					
		% CV	4.2%	14.2%	4.3%					
AR	RC	18	42.7	4.715	113.4	Category 1	Category I	12.0%	3.0%	3.2%
		19	45.7	4.590	114.5					
		20	53.7	4.440	120.3					
		avg	47.3	4.582	116.1					
		S.D.	5.7	0.138	3.7					
		% CV	12.0%	3.0%	3.2%					

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AU	RC	40	50.7	4.285	114.9	Category 1	Category I	8.0%	11.3%	9.7%
		42	48.7	4.465	115.6					
		44	56.7	5.280	135.9					
		avg	52.0	4.677	122.2					
		S.D.	4.2	0.530	11.9					
		% CV	8.0%	11.3%	9.7%					
AV	RC	48	95.7	6.240	189.3	Category 1	Category I	4.6%	8.0%	4.6%
		49	104.7	6.465	201.6					
		51	101.7	5.530	184.6					
		avg	100.7	6.079	191.8					
		S.D.	4.6	0.488	8.8					
		% CV	4.6%	8.0%	4.6%					
AV	RC	19	91.7	3.965	151.1	Category 1	Category I	16.9%	11.6%	13.6%
		20	126.7	4.810	198.8					
		22	101.7	4.950	175.9					
		avg	106.7	4.575	175.3					
		S.D.	18.0	0.533	23.8					
		% CV	16.9%	11.6%	13.6%					
AX	SO	11	154.3	2.273	188.4	Category 1	Category I	15.1%	10.2%	11.3%
		12	116.7	2.218	150.0					
		13	109.0	2.318	143.8					
		14	121.7	2.073	152.8					
		15	110.7	2.708	151.3					
		avg	122.5	2.318	157.3					
		S.D.	18.5	0.237	17.7					
		% CV	15.1%	10.2%	11.3%					

4901

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
AX	SO	1	124.0	2.264	158.0	Category 1	Category I	4.8%	4.2%	3.7%
		2	133.3	2.379	169.0					
		3	136.3	2.239	169.9					
		4	129.0	2.179	161.7					
		5	121.3	2.404	157.4					
		avg	128.8	2.293	163.2					
		S.D.	6.2	0.095	6.0					
		% CV	4.8%	4.2%	3.7%					
AY	RC	41	129.7	5.405	210.7	Category 1	Category I	4.9%	15.4%	8.0%
		42	117.7	4.535	185.7					
		43	123.7	6.195	216.6					
		avg	123.7	5.379	204.3					
		S.D.	6.0	0.830	16.4					
		% CV	4.9%	15.4%	8.0%					
BE	AC	1	5.0	0.290	9.4	Non-irritant	Category III	40.1%	22.5%	32.5%
		2	11.0	0.318	15.8					
		3	12.3	0.240	15.9					
		4	8.0	0.255	11.8					
		5	16.0	0.412	22.2					
		avg	10.5	0.303	15.0					
		S.D.	4.2	0.068	4.9					
		% CV	40.1%	22.5%	32.5%					
BF	SO	35	48.0	1.140	65.1	Category 2A	Category III	18.6%	20.2%	6.5%
		36	33.0	1.722	58.8					
		37	44.0	1.502	66.5					
		avg	41.7	1.455	63.5					
		S.D.	7.8	0.294	4.1					
		% CV	18.6%	20.2%	6.5%					

4902

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
BJ	AL	11	66.7	0.757	78.1	Non-irritant	Category III	9.6%	32.4%	10.2%
		12	67.0	0.777	78.7					
		13	54.7	1.146	71.9					
		14	61.3	0.676	71.4					
		15	70.3	1.396	91.2					
		avg	64.0	0.950	78.3					
		S.D.	6.1	0.308	8.0					
		% CV	9.6%	32.4%	10.2%					
BJ	AL	7	75.0	1.326	94.9	Non-irritant	Category III	4.2%	9.1%	4.9%
		8	70.5	1.341	90.6					
		9	76.5	1.556	99.8					
		avg	74.0	1.408	95.1					
		S.D.	3.1	0.129	4.6					
		% CV	4.2%	9.1%	4.9%					
BM	SO	32	16.0	0.584	24.8	Non-irritant	Category IV	16.1%	59.4%	12.9%
		36	18.0	0.301	22.5					
		37	13.0	1.065	29.0					
		avg	15.7	0.650	25.4					
		S.D.	2.5	0.386	3.3					
		% CV	16.1%	59.4%	12.9%					
BN	SU	1	10.3	0.325	15.2	Non-irritant	Category IV	24.2%	36.0%	21.8%
		2	9.0	0.147	11.2					
		3	5.7	0.369	11.2					
		4	11.0	0.467	18.0					
		5	7.7	0.306	12.3					
		avg	8.7	0.323	13.6					
		S.D.	2.1	0.116	3.0					
		% CV	24.2%	36.0%	21.8%					

4903

Test Material Code	Formulation Type	Cornea Number	Opacity	Permeability	*In vitro* Score	*In vivo* GHS Cat	*In vivo* EPA Cat	Opacity CV	Perm. CV	*In vitro* Score CV
BQ	SO	20	32.0	0.110	33.7	Non-irritant	Category IV	6.9%	48.9%	6.8%
		22	31.0	0.320	35.8					
		23	28.0	0.214	31.2					
		avg	30.3	0.215	33.6					
		S.D.	2.1	0.105	2.3					
		% CV	6.9%	48.9%	6.8%					
BR	SU	43	15.3	0.832	27.8	Non-irritant	Category IV	60.5%	3.7%	30.7%
		46	14.3	0.822	26.7					
		47	3.3	0.776	15.0					
		avg	11.0	0.810	23.2					
		S.D.	6.7	0.030	7.1					
		% CV	60.5%	3.7%	30.7%					
BS	RC	10	208.5	3.478	260.7	Category 2A	Category III	3.7%	15.0%	4.1%
		11	223.5	3.733	279.5					
		12	211.5	4.608	280.6					
		avg	214.5	3.940	273.6					
		S.D.	7.9	0.593	11.2					
		% CV	3.7%	15.0%	4.1%					
EF	RC	40	27.7	5.157	105.0	Category 2A	Category II	19.7%	13.5%	13.4%
		41	18.7	4.797	90.6					
		42	25.7	6.197	118.6					
		avg	24.0	5.384	104.8					
		S.D.	4.7	0.727	14.0					
		% CV	19.7%	13.5%	13.4%					
EG	AC	26	35.7	2.257	69.5	Category 2A	Category II	11.3%	14.0%	4.1%
		27	41.7	2.232	75.2					
		32	44.7	1.742	70.8					
		avg	40.7	2.077	71.8					
		S.D.	4.6	0.290	2.9					
		% CV	11.3%	14.0%	4.1%					
Mean CV for materials with *in vitro* scores greater than 10								**27.9%**	**24.1%**	**18.3%**

Table 7-28 Distribution of product categories for the within-run reproducibility of the BCOP assay. Some products have repeat tests.

Product Categories	Number of products tested
Surfactants	18
Acids	7
Bases	20
Solvents	12
Oxidizers	18
Other	0
Total	**75**

7.3.1.2 BCOP within-run reproducibility for a wide range of materials

The BCOP within-run variability has been assessed in the Bovine Corneal Opacity and Permeability Test Method Background Review Document prepared by NICEATM (appended to this report). The within-run variability for the single parameter of *in vitro* score is presented in Tables 7-1 to 7-3, 7-7 and 7-9 of that document. It should be noted that in each of the tables the mean %CV is significantly influenced by several CV's of greater than 100 generated by the very low overall scores of very mild materials. For example, in Table 7-1 if the last four CV's generated from the extremely low scores of the mild materials are ignored, the mean %CV falls from 48.3% to 18.6%!

7.3.2 BCOP intralaboratory reproducibility

The BCOP interlaboratory reproducibility can be evaluated based not only on data from studies on antimicrobial cleaning products that were submitted to support this specific BRD, but also on the information contained in the Bovine Corneal Opacity and Permeability Test Method Background Review Document prepared by NICEATM.

7.3.2.1 BCOP intralaboratory reproducibility for antimicrobial cleaning products data

Table 7-29 presents intralaboratory reproducibility data for 5 different anti-microbial cleaning products tested from 2 – 6 times in the same laboratory. It can be seen that the individual %CV's range from 2.6 – 49.2%, and the mean intralaboratory %CV for the 5 materials is 20.3%. The highest CV of 49.2% is the result of two extremely high *in vitro* scores, already well above the proposed cutoff of 75 for EPA I or GHS 1 toxicity categories. Thus we consider the impact of this high CV to be negligible. The distribution of product categories for the intralaboratory reproducibility for antimicrobial cleaning products is shown in Table 7-30.

Table 7-29 Intralaboratory reproducibility for 5 antimicrobial cleaning products. See Table 7-27 for individual cornea scores.

Substance	Formulation Type	Mean *In vitro* Irritancy Score (n = 3-5 corneas)	No. of Exp.	Mean	S.D.	%CV
F	RC	360.5 174.4	2	267.5	131.6	49.2%
H	AL	9.2 11.4 15.2 13 13.8 21.3	6	14.0	4.1	29.6%
AV	RC	191.8 175.3	2	183.6	11.7	6.4%
AX	SO	157.3 163.2	2	160.3	4.2	2.6%
BJ	AL	78.3 95.1	2	86.7	11.9	13.7%
Mean %CV					**20.3%**	
Median %CV					**13.7%**	

Table 7-30 Distribution of product categories for the intralaboratory reproducibility for antimicrobial cleaning products.

Product Categories	Number of products tested
Surfactants	0
Acids	0
Bases	2
Solvents	1
Oxidizers	2
Other	0
Total	**5**

7.3.2.2 BCOP intralaboratory reproducibility for a wide range of materials

The BCOP intralaboratory variability has been assessed in the Bovine Corneal Opacity and Permeability Test Method Background Review Document prepared by NICEATM. The intralaboratory variability for the single parameter of *in vitro* score is presented in Tables 7-4 to 7-6 of that document. The mean %CV's for each of the respective tables are: 12.6%, 14.8% and 14.0%

7.3.3 BCOP interlaboratory reproducibility

The Bovine Corneal Opacity and Permeability Test Method Background Review Document prepared by NICEATM is the best source of interlaboratory reproducibility information. This document presents data from an interlaboratory study by Gautheron (1994) (Table 7-31), Balls et al (1995) (Table 7-33) and Southee (1998) (Table 7-35). These data are very useful since the in vitro BCOP data for all of the anti-microbial cleaning products presented in this document were all conducted in one laboratory (IIVS). Thus interlaboratory variability for the specific BCOP antimicrobial cleaning products data could not be assessed.

All three of the above studies are analyzed by %CV of the mean results of the participating laboratories. This is a reasonable way of assessing variability when the results can vary over wide ranges, *e.g.,* as it does with the Cytosensor assay, but it can be extremely biased when used to characterize assays which are constrained at the lower end of irritancy by scores which range around zero. Fluctuations in these scores which are meaningless relative to the entire scoring scale (approximately 500 for the BCOP assay) result in large CV's which inappropriately influence the overall CV of a study by raising the average CV significantly. For example, BCOP scores of 1, 2, and 4 are all indicative of essentially no toxicity and could be said to actually all represent the same score. However, a calculation of the mean and CV of these three values results in a mean of 2.3 and a CV of **65%**! That same variation of one to three units at higher irritancy, *e.g.,* scores of 150, 151, and 154, results in a mean of 152.3 and a CV of **1.0%**! To get an accurate measurement of the true variability of scores, one should scan the range of scores that are being considered and put more weight on the CV's that appear at mid-range than on the scores at the low end of the scale.

The Gautheron study reported data from 11-12 labs. The %CV's were consistently around 30% – 50% (median CV = 46.9%); however, the mean CV was considerably higher at 167%. Inspection of the table shows that this high value is mainly the result of the CV's from mean values which are <5. The distribution of product categories for the interlaboratory reproducibility for the Gautheron study is shown in Table 7-32

Table 7-31 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Gautheron *et al.* (1994)[1]

Substance	Formulation Type	Mean Irritancy Score	No. of Labs	%CV	Prediction
2-Ethoxyethanol	SO	91.3	12	16.5	Severe
2,4-Pentanedione	SO	59.8	12	24	Severe
Allyl alcohol	SO	156	12	27	Severe
Imidazole		87.9	12	28.5	Severe
Furan		56	12	29.4	Severe
Benzethonium chloride	SU	133.9	11	31.7	Severe
Butyrolactone		45.6	12	32.2	Moderate
Cyclohexanone	SO	105.6	11	33.3	Severe
2-Methoxyethanol	SO	63.5	11	33.6	Severe
Laurylsulfobetaine	SU	80.6	11	34	Severe
Ethyl acetoacetate		31.8	11	34.9	Moderate
Gluconolactone		76.6	11	35	Severe
Methylisobutyl ketone	SO	19.9	11	36	Mild
Pyridine	SO	112.8	11	38.4	Severe
Ethanol	SO	60.7	11	39.1	Severe
3-Glycidoxypropyltrimethoxysilane		16.6	12	40	Moderate
N-Lauroylsarcosine, sodium salt	SU	50	11	41.7	Moderate
Octanol	SO	47.4	11	41.7	Moderate
Deoxycholic acid, sodium salt	SU	93.5	12	43	Severe
2-Aminophenol		7	12	43.5	Mild
Hexadecyltrimethylammonium bromide		66.4	11	45.2	Severe
1-Phenyl-3-pyrazolidone		12.9	12	46.5	Mild
Dibenzoyl-L-tartaric acid		120.5	11	46.8	Severe
Dimethyl sulfoxide	SO	11.4	11	46.9	Mild
1-Nitropropane	SO	7.6	12	46.9	Mild
1 ,2,4-Trimethylbenzene		16.1	12	47	Mild
Propyl-4-hydroxybenzoate		7.9	11	48	Mild
Promethazine hydrochloride		112.4	11	49.3	Severe
1 ,2,3-Trichloropropane	SO	47.5	11	50.3	Moderate
Diacetone alcohol	SO	53.5	11	50.8	Moderate
Methanol	SO	84.2	11	55.7	Severe
2,4-Dichloro-5-sulfamoylbenzoic acid		26.3	12	58.5	Moderate
Sodium oxalate		4.8	12	66	Mild
Quinacrine		31.1	11	74.8	Moderate
Petroleum ether	SO	5.5	12	75.4	Mild
Dimethylbiguanide		2.9	11	82	Mild
Magnesium carbonate		3	11	83	Mild
Triethanolamine	SO	2.2	11	101.5	Mild

Substance	Formulation Type	Mean Irritancy Score	No. of Labs	%CV	Prediction
Aluminum hydroxide		6.8	12	107	Mild
Tetraaminopyrimidine sulfate		6	11	107	Mild
Hexane	SO	1.4	12	143	Mild
Iminodibenzyl		2.4	11	177.5	Mild
2-Mercaptopyrimidine		-1.25	12	208	Mild
Triton X-155	SU	0.55	11	276	Mild
DL-Glutamic acid		0.58	12	330.6	Mild
Anthracene		-0.33	12	430	Mild
Betaine monohydrate		0.92	12	432	Mild
MYRJ-45	SU	-0.18	11	962	Mild
EDTA di-potassium salt		-0.33	12	1009	Mild
BRIJ-35	SU	-0.09	11	1280	Mild
Phenylbutazone		-0.17	12	1325	Mild
Mean CV(%)					**167.6 (all substances)** **84 (excluding MYRJ-45, EDTA, BRIJ-35, phenylbutazone)**
Median CV(%)					**46.9**

[1] *Substances organized by increasing %CV.*

Table 7-32 Distribution (estimated) of product categories for the interlaboratory reproducibility for the Gautheron study.

Product Categories	Number of products tested
Surfactants	7
Acids	Unknown
Bases	Unknown
Solvents	16
Oxidizers	Unknown
Other	Unknown
Total	**23**

The Balls *et al.* study reported data from 5 labs. The %CV's were consistently around 25%-35% (median CV = 30.6%); however, the mean CV was considerably higher at 125%. Again, inspection of the table shows that this very high value is mainly the result of the CV's from mean values which are <5. The distribution of product categories for the interlaboratory reproducibility for the Balls study is shown in Table 7-34.

Table 7-33 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Balls *et al.* (1995)

Substance	Formulation Type	Mean Irritancy Score	No. of Labs	%CV	Prediction
1 -Naphthalene acetic acid, Na salt		149.2	5	7.6	Severe
Benzalkonium chloride (10%)	SU	136.5	5	10.9	Severe
Sodium hydroxide (1%)	AL	150	5	12.3	Severe
Cetylpyridinium bromide (6%)	SU	71.2	5	12.7	Severe
Acetone	SO	123	5	14	Severe
Imidazole		112.7	5	14.5	Severe
Benzalkonium chloride (5%)	SU	128.5	5	15.6	Severe
Methyl acetate	SO	54.9	5	17.4	Moderate
Sodium hydroxide (10%)	AL	271.9	5	17.6	Severe
Toluene	SO	35.6	5	18.1	Moderate
Chlorhexidine		114	5	18.3	Severe
Trichloroacetic acid (3 0%)	AC	264	5	18.7	Severe
Dibenzyl phosphate	SO	378	5	18.8	Severe
2,2-Dimethylbutanoic acid	AC	111.9	5	19.5	Severe
Pyridine	SO	148	5	20.1	Severe
Promethazine hydrochloride		121.4	5	20.4	Severe
Trichloroacetic acid (3%)	AC	75.9	5	21.1	Severe
Benzalkonium chloride (1 %)	SU	88.8	5	21.7	Severe
Parafluoraniline		30.4	5	21.7	Moderate
Methyl ethyl ketone	SO	70.4	5	22.6	Severe
4-Carboxybenzaldehyde		78.3	5	24	Severe
Ethanol	SO	70.6	5	24.1	Severe
Cetylpyridinium bromide (10%)	SU	72	5	24.2	Severe
Triton X-100 (5 %)	SU	78.3	5	24.2	Severe
Triton X-100 (10 %)	SU	70.3	5	25.3	Severe
Isobutanol	SO	56	5	26.1	Severe
n-Hexanol	SO	61.9	5	27	Severe
Sodium lauryl sulfate (15 %)	SU	63.3	5	28	Severe
Cyclohexanol	SO	60.1	5	28.5	Severe
2,6-Dichlorobenzoyl chloride		10.4	5	30.6	Mild
Sodium lauryl sulfate (3 %)	SU	25.8	5	30.9	Mild
Isopropanol	SO	57.9	5	31.3	Severe
Sodium perborate		97	5	35.8	Severe
Methyl isobutyl ketone	SO	12.6	5	36	Mild
1-Naphthalene acetic acid		78.1	5	37.4	Severe
Butyl acetate	SO	34.6	5	38.4	Moderate
Methyl cyanoacetate		12.2	5	39.2	Mild
Ethyl acetate	SO	32	5	40.5	Moderate
Potassium cyanate		15	5	40.9	Mild
2,5-Dimethylhexanediol	SO	20.8	5	41.6	Mild
Benzoyl-L-tartaric acid		169.6	5	43	Severe
gamma-Butyrolactone	SO	60.7	5	45	Severe
Tetraaminopyrimidine sulfate		15.1	5	46.3	Mild
Methylcyclopentane	SO	2.8	5	47.8	Mild
2-Ethyl-1-hexanol	SO	39.8	5	48.2	Moderate
Cetylpyridinium bromide (0.1%)	SU	9.2	5	51.4	Mild

Substance		Mean Irritancy Score	No. of Labs	%CV	Prediction
Maneb		40.5	5	58.3	Moderate
n-Octanol	SO	40.9	5	58.8	Moderate
Ethyl-2-methylacetoacetate		14.4	5	65.3	Mild
Ethyl trimethyl acetate	SO	17.8	5	66.3	Mild
Ammonium nitrate		9.8	5	69.7	Mild
L-Aspartic acid		1.3	5	73.6	Mild
Captan 90 concentrate		43.8	5	75.8	Moderate
Quinacrine		1.6	5	76.9	Mild
Fomesafen		60.7	5	89.4	Severe
Sodium oxalate		14	5	143	Mild
Polyethylene glycol 400	SU	1.1	5	145	Mild
Glycerol	SO	0.26	5	712	Mild
Tween 20	SU	-0.04	5	4511	Mild
Mean %CV					**125 (all test substances)** **50 (excluding Tween 20)**
Median %CV					**30.6**

[1]*Substances organized by increasing %CV.*

Table 7-34 Distribution of product categories (estimated) for the interlaboratory reproducibility for the Balls study.

Product Categories	Number of products tested
Surfactants	12
Acids	1
Bases	2
Solvents	21
Oxidizers	Unknown
Other	Unknown
Total	**36**

The Southee *et al.* study reported data from 3 labs. The% CV's were consistently around 15% - 25%% (median CV = 22.8%); however, the mean CV was higher at 32%. Again, inspection of the table shows that this higher value is mainly the result of the CV's from mean values which are <5. The distribution of product categories for the interlaboratory reproducibility for the Southee study is shown for Table 7-36.

Table 7-35 Coefficient of Variation Analysis of the Interlaboratory Variability of the BCOP Test Method for Southee (1998)

Substance	Formulation Type	Mean Irritancy Score	No. of Labs	%CV	Prediction
Butyl cellosolve	SO	100.9	3	7.5	Severe
Benzalkonium chloride	SU	160	3	8.5	Severe
NaOH (10%)	AL	226	3	8.6	Severe
Imidazole		136.9	3	9.1	Severe
4-Carboxybenzaldehyde		46.7	3	9.5	Moderate
Parafluoroaniline		32.1	3	19.1	Moderate
Methyl ethyl ketone	SO	82.5	3	21.6	Severe
Ethanol	SO	48.7	3	22.1	Moderate
Ammonium nitrate		5.03	3	23.4	Mild
Hexadecyltrimethylammonium bromide (10%)		29.3	3	27.1	Moderate
Glycerol	SO	0.72	3	33.5	Mild
Propyl-4-hydroxybenzoate		6.9	3	37.7	Mild
Triton X-100 (5%)	SU	3.3	3	44.8	Mild
Sodium lauryl sulfate (15%)	SU	9.7	3	57.1	Mild
Tween 20	SU	0.23	3	79.8	Mild
Sodium oxalate		3.6	3	108.8	Mild
Mean %CV					**32.4**
Median %CV					**22.8**

[1]*Substances organized by increasing %CV*

Table 7-36 Distribution of product categories (estimated) for the interlaboratory reproducibility for the Southee study.

Product Categories	Number of products tested
Surfactants	4
Acids	Unknown
Bases	1
Solvents	4
Oxidizers	Unknown
Other	Unknown
Total	**9**

Additional information on interlaboratory reproducibility is given in Section 7 of the NICEATM BRD on the BCOP assay.

8 Test Method Data Quality

8.1 Adherence to National and International GLP Guidelines

ICCVAM has suggested (ICCVAM 1997) that all data submitted in support of a new method be generated by methodologies that adhere to national or international GLP guidelines. It could not be ascertained that all of the *in vitro* data contained in this BRD were generated under full GLP compliance, but where it could be ascertained, that information is contained in the spreadsheets that form the database from which this BRD was generated. All of the new *in vitro* data that were generated during the course of constructing this BRD were conducted with full GLP compliance.

8.2 Data Quality Audits

No data quality audits were conducted for the purpose of this BRD. Complete GLP audits were, of course, conducted for the studies included in this BRD which are identified as having been conducted with full GLP compliance.

8.3 Impact of Deviation from GLP Guidelines

The data were not evaluated for the effect of any GLP deviations that may have been noted. However, *in vitro* data were accompanied by information that Criteria for a Valid Test listed in the protocol had been fulfilled during the study.

8.4 Availability of Laboratory Notebooks or Other Records

Study notebooks, final reports, and other background documents are available for the majority of *in vitro* studies reported here. These documents have not been included with this BRD, but they will be available in a confidential form for inspection upon the request of NICEATM or the EPA. Companies who submitted data for this BRD did so with the understanding that their identities would not be linked to any of the tested materials. Thus company identifiers will be removed from any study notebooks or final reports which are requested by NICEATM or the EPA for audit.

9 Other Scientific Reports and Reviews

The three *in vitro* methodologies (Cytosensor, EpiOcular, and BCOP assay) that are addressed in this BRD have been the subject of three individual BRD's. Since two are still under review, only the BCOP BRD is appended to this BRD.

The Cytosensor BRD was created under contract for ECVAM by IIVS. It has undergone an independent data audit and has been reviewed for scientific content by an independent management team designated by ECVAM. The final review for validity of the method has not been completed but is underway.

A BRD for the EpiOcular model has been created under contract to the Colgate-Palmolive company by IIVS. It has been submitted to ECVAM and has undergone a preliminary review by the ocular toxicology task force. Modifications and additions have been made to the document at the request at the task force, and it was resubmitted to ECVAM in December 2007.

A BRD for the BCOP assay was created by NICEATM as part of their program to identify the "Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants". It has been reviewed and modified and exists on the internet at

http://iccvam.niehs.nih.gov/methods/ocutox/ivocutox/ocu_brd_bcop.htm.

10 Animal Welfare Considerations

10.1 How the proposed non-animal testing strategy will refine, reduce or replace animal use for the purpose of toxicity labeling of anti-microbial cleaning products

Currently the EPA Pesticides Program requires a rabbit eye irritation test to determine the correct toxicity labeling category for ocular irritation. This test requires between one (if the material is shown to be corrosive or severe) and three (to determine less severe categories) rabbits for each product submitted for registration. The testing strategy proposed in this BRD will completely replace the use of laboratory animals in this registration process. Thus, no animals will be subjected to pain and suffering for the purpose of determining the EPA labeling category for eye irritation.

In terms of overall animal use, the BCOP assay does use tissue from animals (cattle), but these animals have already been slaughtered for the purpose of food production at the time that the ocular tissue is obtained. The cattle undergo no additional pain or suffering during the harvesting of the corneal tissue. In addition, this tissue is normally discarded and would end up being wasted if it was not used to prevent suffering to live animals.

Neither the Cytosensor method nor the EpiOcular method use live animals. A long established mouse cell line is used in the Cytosensor assay, but no new animals have to be sacrificed to conduct the assay. The EpiOcular assay utilizes only human tissue, and thus completely avoids the use of non-human animals.

Thus, no animals are harmed or experience pain as a result of conducting any of the three *in vitro* ocular irritation assays proposed in this testing strategy.

11 Practical Considerations

11.1 Use by industry

All of the companies involved in this validation effort conduct safety evaluations of new cleaning products in a similar fashion to that shown in Figure 11-1. However, the toxicity evaluation portion may be conducted differently depending on the individual company's specific product types and corporate experience and expertise. The specific *in vitro* test or tests to be used are chosen with knowledge of the historical performance of specific types of product chemistry with specific *in vitro* tests. The test data that result may then be compared to previous product test results contained in an historical database.

Since no one company has managed to develop comprehensive experience with the *in vitro* ocular activities of the entire range of anti-microbial cleaning products (because each company manufactures only specific product lines), we have attempted in this BRD to combine the experience and knowledge of all the companies. Thus each company's specific experience with its product line has been combined with that of others to produce a broad, generalized approach which covers the range of product types which exist in today's market place and are anticipated to be marketed in the reasonable future.

Most of these companies have spent a number of years developing the data we have presented here in the course of creating a safety evaluation approach which protects consumers without the use of whole animal studies.

Figure 11-1 Process of safety evaluations

Each of the tests described in this BRD has a long history of use by industry. The history of use of the BCOP assay has been documented in detail by NICEATM in the Background Review Document (BRD) "Current Status of *In Vitro* Test Methods for Identifying Ocular Corrosives and Severe Irritants: Bovine Corneal Opacity and Permeability Test Method" (NIEHS 2006). The BCOP assay is used as an in-house screen to assess potential ocular irritation of a wide range of substances resulting from accidental exposure in the workplace or home (Sina 1994; Swanson, Lake et al. 1995; Casterton, Potts et al. 1996; Chamberlain, Gad et al. 1997; Harbell and Curren 1998; Cater, Nusair et al. 2002; Cuellar, Lloyd et al. 2003; Bailey, Freeman et al. 2004). A secondary application has been the use of the

assay for product development. By comparing new formulations with marketed materials of similar composition, manufacturers can evaluate the relative irritancy as part of an initial safety screening. The BRD cites specific instances to illustrate the applications of the assay for workplace safety (Chamberlain, Gad et al. 1997), product safety, and/or product development (Cuellar N and Swanson J, personal communications). In both instances, *in vivo* confirmatory testing is reduced or may not be necessary. Details of the use of this assay can be found in Section 2.2.3.

The Cytosensor microphysiometer assay has been used by companies since its first introduction in the early 1990's (Bruner, Miller et al. 1991) to assess eye irritancy potential of liquid or water soluble ingredients or formulations. This information is sometimes combined with other available information in integrative toxicological evaluation, and final safety decisions are made based on this information. Details can be found in the background description Section 2.2.1.

The EpiOcular model is also used by industry as an *in vitro* assay to assess eye irritation potential (Ghassemi, Osborne et al. 1997; Stern, Klausner et al. 1998). As with the other two models, this information is sometimes combined with other available information in integrative toxicological evaluation, and final safety decisions are made based on this information. Details can be found in the background description Section 2.2.2.

11.2 Ease of transferability

ECVAM has recently (Hartung, Bremer et al. 2004) discussed transferability of tests. In their manuscript describing a modular approach to validation, they state that transferability "should demonstrate that the test can be successfully repeated in a laboratory different from the one which has developed or which was involved in the optimization of the test". Ease of transferability is supported by evaluating interlaboratory reproducibility of the assay system. Details of interlaboratory reproducibility for each of the three *in vitro* methods addressed in this BRD are given in Section 7 of this BRD and in even greater detail in the accessory BRD that is appended to this report.

11.2.1 Facilities and major fixed equipment for the Cytosensor test method

The major fixed equipment for the Cytosensor test method is the instrument itself. When purchased new the instrument was quite expensive (>$100,000), but as of Summer 2007 the instrument is no longer available from its former manufacturer, Molecular Devices, Inc. (Menlo Park, CA). In addition, Molecular Devices has stated that they will be able to sell disposable supplies for the machine only until their current supply lasts. At this time (Summer 2008), we have discovered that they have provided at least one user with the name of their third-party contractor(s) who manufactures the disposables for them. This user has found that purchasing the

disposable supplies directly from the original manufacturer (not Molecular Devices) is possible, and he has shared the purchasing information with IIVS. None the less, in anticipation of the dropping of support for the instrument, IIVS has purchased a supply of disposables which should last for at least two years, and they have also obtained repair parts that are also likely sufficient to support the instrument for more than two years. Thus it is likely that testing can continue with the CM for a number of years.

More information on additional standard laboratory equipment needed to support this assay is given in detail in the Cytosensor BRD prepared for ECVAM which will be available after it has had a final review by ESAC.

11.2.2 Facilities and major fixed equipment for the EpiOcular test method

A general purpose tissue culture facility is required for the use of the EpiOcular model. There should be provisions for handling the cultures in a sterile environment as well as facilities for appropriately containing any toxic test materials that might be utilized in the test.

Major equipment would include a Class II Type A or B tissue culture hood, 37°C humidified incubator, and an inverted microscope. A 96-well plate reader is highly desirable but not mandatory since a small spectrophotometer could also be used.

More information on additional standard laboratory equipment needed to support this assay is given in detail in the EpiOcular BRD prepared for ECVAM which will be attached to this BRD after it has had a final review by ESAC.

11.2.3 Facilities and major fixed equipment for the BCOP test method

The main facility requirements for the BCOP assay can be found in most any standard biology laboratory. Sterile handling of the tissue is not an absolute requirement and most experiments can be conducted on the bench top. Proper containment is, of course, needed anytime that extremely toxic materials are tested. One major piece of equipment required is the opacitometer which can be obtained from Stag Bio (Clermont, France). The price of the opacitometer has risen significantly over the years and is now quoted at ~$9000 per unit.

More information on additional standard laboratory equipment needed to support this assay is given in detail in the BCOP BRD (ICCVAM 2006) prepared by NICEATM which will be attached to this BRD after it has had a final review by ESAC.

11.3 Training required

Training in standard *in vitro* techniques and laboratory procedures is required for all of the three assays in order to assure that the assay is run correctly. Since it is likely that toxic materials may be tested in the assays, laboratory safety training should also be required before a technician is allowed to conduct any of the assays.

11.3.1 Required level of training and expertise needed to conduct the Cytosensor assay

Two areas of training are especially important for the conduct of the Cytosensor assay. The first is general tissue culture technique needed to culture the L929 cells (or other cell lines) which are used as the target cells in the assay. The second is specific operation of the Cytosensor itself. Although many of the functions of the machine are programmed to occur automatically through the supplied CytoSoft program, the technician still needs to learn how to program the general parameters of each run into the controlling computer. This training is not arduous, but should be continued until the technician can reproducibly test 3 to 5 compatible materials with the Cytosensor such the values for the materials approach the historic mean for those materials tested in that laboratory.

11.3.2 Required level of training and expertise needed to conduct the EpiOcular assay

The techniques involved with the EpiOcular methodology are fairly standard for those trained within an *in vitro* toxicology laboratory. No specific expertise outside of that commonly used for tissue culture and toxic material handling is required.

Training for this specific method is required and is assisted by developing a detailed laboratory workbook that outlines the procedures and the data that need to be recorded at each step.

In the IIVS laboratory, each technician is required to demonstrate acceptable performance for the testing of five standard surfactant materials whose toxicities are well established.

11.3.3 Required level of training and expertise needed to conduct the BCOP assay

The following discussion of training for the BCOP assay is abstracted from the NICEATM-prepared BCOP BRD appended to this report.

"A training period of between two to three months is usually required for a technician with general laboratory skills to proficiently conduct all aspects of the standard BCOP assay with reasonably little supervision. The individual would need basic laboratory skills including

- Aseptic technique,
- Knowledge and training in the preparation of dilutions,
- Training in the use of an analytical balance,
- Proficiency in the use of single channel pipettes
- Calibration and use of the spectrophotometer

Specific laboratory skills would include

- Identification of a bovine cornea free of corneal defects
- Excising the cornea from the bovine eye
- Mounting the cornea in a corneal holder without damaging the epithelium or endothelium
- Addition of media without air bubbles to the posterior and anterior chamber
- Examination of mounted corneas for defects
- Addition of test material to the corneal chamber
- Washing the cornea (closed and open chamber) without inducing mechanical damage
- Calibration and use of the opacitometer"

"Evaluation of a technician for proficiency in the assay is based upon the successful performance of the assay using positive and negative controls.

The histopathological evaluation of the corneas requires skills in the preparation and the evaluation of corneal tissue. Fixed corneas should be trimmed, embedded and stained by a qualified histology laboratory. Proficiency in the evaluation of the slides requires a training period of up to six months and is dependent on the experience of the individual."

11.4 Cost Considerations

A GLP-compliant CM assay conducted at IIVS is $2,050 (minimum of 2 test materials). Five or more materials run concurrently is $1,375 per test material. These prices are currently used at IIVS. We know of no other commercial sources for the CM assay.

A GLP-compliant EO assay (range-finding plus definitive assay; positive and negative control; duplicate tissues) conducted at IIVS is $3,700 for a single test article. Five or more materials run concurrently is $2,750 per test material. A second laboratory, MB Research Laboratories (Spinnerstown, PA), charges $2,200 per test

article for two replicates at three time points and charges $3,225 for four time points for one test article.

The price for a GLP compliant BCOP assay at IIVS is approximately $1,850 for a single test substance, including positive and negative controls. Histopathology can be performed on corneas from that same study for an additional $4,750. Costs per test substance can be reduced considerably with the performance of multiple materials run concurrently. For example, a single material tested in the BCOP assay plus histology would be approximately $6,600, two materials run concurrently would be approximately $3,900/test material, three materials run concurrently would be approximately $3,480/test material and four materials run concurrently would be approximately $3,300/test article. A second laboratory, MB Research Laboratories (Spinnerstown, PA), charges $1,000 per test article with no histology and $1,900 per test article with histology.

The price for a GLP compliant *in vivo* Driaze test ranges from approximately $1160 to $14,500 depending on the lab and the number of days the animals remained on study.

Unlike *in vivo* testing, *in vitro* testing lends itself to significant economies of scale when evaluating multiple materials concurrently.

11.5 Time Considerations

Timing for each of the three assays varies and is described for each below. These times should be compared with a typical *in vivo* rabbit eye test which would require a minimum of one to three days, although the assay must be extended up to 21 days if certain lesions don’t clear.

11.5.1 Timing for Cytosensor test method

The Cytosensor assay can actually be conducted in a single day, including multiple runs of the test material. Completion of the final report would then take several more days.

11.5.2 Timing for EpiOcular test method

The EpiOcular test generally takes one and one-half to two days in the laboratory to complete. A two week lead time is usually required to obtain the EpiOcular tissue from its manufacturer, MatTek Corporation. Again completion of the final report would take several more days.

11.5.3 Timing for BCOP test method

The standard BCOP assay can be completed in the laboratory in one day (an extended day may be necessary for certain protocol modifications). Completion of the final report would take several more days.

If histology is required, *e.g.,* if the BCOP score was <75, but >25, then turnaround time would be considerably extended. Currently at IIVS it can take several weeks to have the tissue processed and then more time to have the slides read by a pathologist.

Total time required for the assay if histology was require would be approximately four weeks.

12 References

Araki, K., Y. Ohahsi, S. Knoshita, K. Hayahsi, Y. Kuwayama and Y. Tano (1994). "Epithelial Wound Healing in the Denervated Cornea." Current Eye Research **13**: 203-211.

Bailey, P., J. Freeman, R. Phillips and J. Merrill (2004). "Validation of the BCOP Assay as a Predictor of Ocular Irritation of Various Petrochemical Products." Society of Toxicology 2004 Meeting.

Baldwin, H. A., T. O. McDonald and C. H. Beasley (1973). "Slit examination of experimental animals eyes. II Grading scales and photographic evaluation of induced pathological conditions." J Society of Cosmetic Chemistry **24**: 181-195.

Balls, M., P. A. Botham, L. H. Bruner and H. Spielmann (1995). "The EC/HO International Validation Study on Alternatives to the Draize Eye Irritation Test." Toxicology In Vitro **9**(6): 871-929.

Beckley, J. H. (1965). "Comparative Eye Testing: Man vs. Animal." Toxicol Appl Pharmacol **7**: 93-101.

Beckley, J. H. (1969). "Use of the Rhesus Monkey for Predicting Human Response to Eye Irritants." Toxicol Appl Pharmacol **15**: 1-9.

Bell, G. H., D. Emslie-Smith and C. R. Paterson (1976). Chapter 41: Special Senses. Edinburgh.

Botham, P., R. Osborne, K. Atkinson, G. Carr, M. Cottin and R. G. Van Buskirk (1997). "IRAG Working Group 3. Cell Function-based Assays. Interagency Regulatory Alternatives Group." Food and Chemical Toxicology **35**(1): 67-77.

Brantom, P. G., L. H. Bruner, M. Chamberlain, O. De Silva, J. Dupuis, L. K. Earl, 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.

Bruner, L. H., G. J. Carr, R. D. Curren and M. Chamberlain (1998). "Validation of Alternative Methods for Toxicity Testing." Environ Health Perspect **106 Suppl 2**: 477-84.

Bruner, L. H. and K. A. Kohrman (1993). The Low Volume Eye Test as a Refinement to the Draize Eye Irritation Test. Bethesda, Scientists Centre for Animal Welfare.

Bruner, L. H., K. M. Miller, J. C. Owicki, J. W. Parce and V. C. Muir (1991). "Testing Ocular Irritancy In Vitro with the Silicon Microphysiometer." Toxicology In Vitro **5**: 277-284.

Burdick, J., J. Merrill, T. Spangler, G. Moyer and J. Harbell (2002). "Use of Histological Examination in Bovine Corneal Opacity and Permeability (BCOP) Assay for Assessing the Ocular Irritation Potential of Fragranced Formulations." The Toxicologist **66**(1): 244.

Butler, J. M. and B. R. Hammond (1980). "The Effects of Sensory Denervation on the Response of the Rabbit Eye to prostaglandin E1, bradykinin and substance P." British Journal of Pharmacology **69**: 495-502.

Casterton, P. L., L. Potts and B. Klein (1996). "A Novel Approach to Assessing Eye Irritation Potential Using the Bovine Corneal Opacity and Permeability Assay." J Toxicol-Cut and Ocular Toxicol **15**(2): 147-163.

Cater, K. C., G. Mun, G. Moyer, J. Merrill and J. Harbell (2003). "Exploratory In Vitro Eye Irritation Study of Marketed Alkaline Dry Laundry Detergents by BCOP Assay and pH/reserve Alkalinity (RA) Parameters." The Toxicologist **72**: 220.

Cater, K. C., T. Nusair, J. C. Merrill and J. W. Harbell (2002). "Exploratory In Vitro Eye Irritation Study of Marketed Alkaline Liquid Laundry Detergents by BCOP Assay and pH/reserve Alkalinity (RA) Parameters." The Toxicologist **66**: 244.

Cater, K. C., H. A. Raabe, G. C. Mun and J. W. Harbell (2001). "Corporate Validation Program for Predicting Eye Irritation of Surfactant Formulations." The Toxicologist **60**(1): 99.

Chamberlain, M., S. C. Gad, P. Gautheron and M. K. Prinsen (1997). "IRAG Working Group 1. Organotypic Models for the Assessment/prediction of Ocular Irritation. Interagency Regulatory Alternatives Group." Food Chem Toxicol **35**(1): 23-37.

Chan, P. K. and A. W. Hayes (1994). "Acute Toxicity and Eye Irritancy." In Principles and Methods of Toxciology, 3rd Edition (ed Hayes A.W.): 579-647.

Choksi, N., J. Haseman, D. Allen, R. Tice and W. Stokes (2005). Estimated Likelihood for Under- and Over-Classification for a Sequential Draize Rabbit Eye Test. Society of Toxicology 45th Annual Meeting, San Diego, CA.

Chrai, S., T. F. Patton, A. Mehta and J. R. Robinson (1973). "Lacrimal and Instilled Fluid Dynamics in Rabbit Eyes." J Pharmascent Sci **62**: 1112-1121.

Cooper, K. J., L. K. Earl, J. Harbell and H. Raabe (2001). "Prediction of Ocular Irritancy of Prototype Shampoo Formulations by the Isolated Rabbit Eye (IRE) Test and Bovine Corneal Opacity and Permeability (BCOP) Assay." Toxicol In Vitro **15**(2): 95-103.

Cormier, e. M., J. E. Hunter, W. Billhimer, J. May and M. A. Farage (1995). "Use of Clinical and Consumer Eye Irritation Data to Evaluate the Low-volume Eye Test." J Toxicol-Cut and Ocular Toxicol **14**: 197-205.

Cuellar, N., P. H. Lloyd, J. E. Swanson, J. Merrill, M. Clear, G. Mun, et al. (2003). "Evaluating the Eye Irritancy of Solvents in a Simple Fragrance Mixture with the Bovine Corneal Opacity and Permeability (BCOP) Assay." The Toxicologist **72**(1): 312.

Cuellar, N., P. H. Lloyd, J. E. Swanson, J. Merrill, G. Mun, J. W. Harbell, et al. (2004). "Phase Two: Evaluating the Eye Irritancy of Solvents in a Single Fragrance Mixture with the Bovine Corneal Opacity and Permeability (BCOP) Assay." The Toxicologist **78**(1): 268.

Curren, R., M. G. Evans, H. A. Raabe, R. R. Ruppalt and J. W. Harbell (2000). "An Histopathological Analysis of Damage to Bovine Corneas In Vitro by Selected Ocular Toxicants." The Toxicologist **54**(1): 188.

Daston, G. P. and F. E. Freeberg (1991). Chapter 16 Ocular Irritation Testing. Columbus, Ohio, CRC Press.

Dearman, R. J., M. Cumberbatch and I. Kimber (2003). "Cutaneous Cytokine Expression: Induction by Chemical Allergen and Paracrine Regulation." Journal of Toxicology-Cutaneous and Ocular Toxicology **22**: 69-86.

DeSousa, D. J., A. A. Rouse and W. J. Smolon (1984). "Statistical Consequences of Reducing the Number of Rabbits Utilized in Eye Irritation Testing: Data on 67 Petrochemicals." Toxicol Appl Pharmacol **76**: 234-242.

Draize, J. H., G. Woodard and H. O. Calvery (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.

Draize, J. H., G. Woodward and H. O. Calvery (1944). "Methods for the study of irritation and toxicity of the substances applied to the skin and mucous membranes." J Pharmacol. **82**: 377-390.

Dua, H. S., J. A. P. Gomes and A. Singh (1994). "Corneal Epithelium Wound Healing." British Journal of Ophthamology **78**: 401-408.

ECETOC (1992). "ECETOC Technical Report No 48."

ECETOC (1998). "ECETOC Technical Report No. 48 (2)."

Ehlers, N. (1976). Pharmacology of the Conjunctival Sac. Drugs and Ocular Tissues, Second Meeting of the International Society for Eye Research.

EPA (1998). "Health Effects Test Guideline, OPPTS 870.2400 Acute Eye Irritation. EPA 712-C-98-195." U.S. Environmental Protection Agency: Washington, DC.

EPA (2003). "Label Review Manual. 3rd Edition. EPA 735-B-03-001." U.S. Environmental Protection Agency: Washington, DC.

Eurell, T. E., J. M. Sinn, P. A. Gerding and C. L. Alden (1991). "In Vitro Evaluation of Ocular Irritatns Using Corneal Protein Profiles." Toxicology and Applied Pharmacology **108**: 374-378.

Evans, M. G. (1998). "Condiserations for Histopathological Examination of Bovine Corneal Tissue." In Vitro & Molecular Toxicology **11**: 335-337.

Farquhar, M. G. and G. E. Palade (1963). "Junctional complexes in Various Epithelia. ." J Cell Biology **17**: 375-412.

Feder, P. I., R. A. Lordo, D. M. B. Dipasquale, M. Chudkowski, J. L. Demetrulias, K. L. Hintze, et al. (1991). "The CTFA Evaluation of Alternatives Program: An Evaluation of *In Vitro* Alternatives to the Draize Primary Eye Irritation Test (Phase I) Hydro-alcoholic Formulations; (Part 1) Statistical Methods." In Vitro Toxicology **4**(4): 231-246.

Freeberg, F. E., J. F. Griffith, R. D. Bruce and P. H. S. Bay (1984). "Correlation of Animal Test Methods with Human Experience for Household Products." J. Toxic. Cut. & Ocular Toxic **1**(3): 53-64.

Freeberg, F. E., D. T. Hooker and J. F. Griffith (1986). "Correlation of Animal Test Methods with Human Experience for Household Products: An Update." J Toxicol-Cut and Ocular Toxicol **5**(2): 115-123.

Freeberg, F. E., G. A. Nixon, P. J. Reer, J. E. Weaver, R. D. Bruce, J. F. Griffith, et al. (1986). "Human and Rabbit Eye Responses to Chemical Insult." Fund. Appl. Toxic. **7**: 626-634.

Gautheron, P., J. Giroux, M. Cottin, A. Audegond, A. Morilla, L. Mayordomo-Blanco, et al. (1994). "Interlaboratory Assessment of the Bovine Corneal Opacity and Permeability (BCOP) Assay." Toxicology In Vitro **8**(3): 381-392.

Gettings, S. D., L. C. Dipasquale, D. M. Bagley, P. L. Casterton, M. Chudkowski, R. D. Curren, et al. (1994). "The CTFA Evaluation of Aternatives Program: An Evaluation of *In Vitro* Alternatives to the Draize Primary Eye Irritation Test. (Phase II) Oil/Water Emulsions." Fund. Chem. Toxic. **32**(10): 943-976.

Gettings, S. D., R. A. Lordo, K. L. Hintze, D. M. Bagley, P. L. Casterton, M. Chudkowski, et al. (1996). "The CTFA Evaluation of Alternatives Program: an evaluation of in vitro alternatives to the Draize primary eye irritation test. (Phase III) surfactant-based formulations." Food Chem Toxicol **34**(1): 79-117.

Ghassemi, A., R. Osborne, K. Kohrman, M. Roddy, Her and B. Kanengiser (1997). "Demonstrating the Ocular Safety of an Eye Cosmetic Product Using Alternatives to Animal Eye Irritation Tests." The Toxicologist: Fundamental and Applied Toxicology Supplement **36**(1, Part 2): 43.

Ghassemi, A., L. H. Sauers, L. Bruner, P. J. Reer and R. H. Hall (1993). "Demonstrating the Human Safety of a New Household Cleaning (HSC) Products Using Alternatives to the Draize Eye Irritation Test." Presentation made at the US Society of Toxicology Meeting.

Gran, B. P., J. E. Swanson, J. Merrill and J. Harbell (2003). "Evaluation the Irritancy Potential of Sodium Percarbonate: A Case Study Using the Bovine Corneal Opacity and Permeability (BCOP) Assay." The Toxicologist **72**: 220.

Griffith, J. F., G. A. Nixon, R. D. Bruce, P. J. Reer and E. Bannan (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.

Hackett, R. B. and T. O. McDonald (1994). "Mechanisms of Ocular Response to Irritants." Dermatotoxicology, 5th Editions, (ed Marzulli F.N, Maibach H.I.): 299-306.

Harbell, J., S. W. Koontz, R. W. Lewis, D. Lovell and D. Acosta (1997). "IRAG Working Group 4: Cell Cytotoxicity Assays." Food and Chemical Toxicology **35**: 79-126.

Harbell, J. W. and R. D. Curren (1998). "The Bovine Corneal Opacity and Permeability Assay: Observation on Assay Performance." In Vitro & Molecular Toxicology **11(4)**: 337-341.

Harbell, J. W., R. Osborne, G. J. Carr and A. Peterson (1999). "Assessment of the Cytosensor™ Microphysiometer Assay in the COLIPA In Vitro Eye Irritation Validation Study." Toxicology In Vitro **13**: 313-323.

Hartung, T., S. Bremer, S. Casati, S. Coecke, R. Corvi, S. Fortaner, et al. (2004). "A Modular Approach to the ECVAM Principles on Test Validity " Atern Lab Anim **32**: 467-472.

Hogan, M. J. and L. E. Zimmerman (1962). "Ophthalmic Pathology: An Atlas and Textbook, 2nd edition." Philadelphia: W.B. Saunders.

Hubert, F. (1992). "The Eye (Rabbit/Human): Parameters to be Measured in the Field of Ocular Irritation." ATLA **20**: 476-479.

ICCVAM (1997). "Validation and Regulatory Acceptance of Toxicological Test Methods: A Report of the ad hoc Interagency Coordinating committee on the Validation of Alternative Methods." NIH Publication No.: 97-3981. Research Triangle Park: National Toxicology Program.

ICCVAM (2006). *In Vitro* Test Methods for Detecting Occular Corrosives and Severe Irritants.

Jacobs, G. A. and M. A. Martens (1990). "Quantification of Eye Irritation Based Upon In Vitro Changes of Corneal Thickness." ATLA **17**: 255-262.

Jester, J. V., L. Li, A. Molai and J. K. Maurer (2001). "Extent of initial corneal injury as a basis for alternative eye irritation tests." Toxicol In Vitro **15**(2): 115-30.

Jester, J. V., W. M. Petroll, J. Bean, R. D. Parker, G. J. Carr, H. D. Cavanagh, et al. (1998). "Area and depth of surfactant-induced corneal injury predicts extent of subsequent ocular responses." Invest Ophthalmol Vis Sci **39**(13): 2610-25.

Jones, P. A., E. Budynsky, K. J. Cooper, D. Decker, H. A. Griffiths and J. H. Fentem (2001). "Comparative Evaluation of Five In Vitro Tests for Assessing the Eye Irritation Potential of Hair-care Products." ATLA **29**: 669-692.

Karson, C. N., K. F. Berman, E. F. Donnelly, W. B. Mendelson, J. E. Kleinman and R. J. Wyatt (1981). "Speaking, Thinking, and Blinking." Psychiatry Research **5**: 243-246.

Katahira, J. H., H. Sugiyama, N. Inoue, Y. Horiguchi, M. Matsuda and N. Sugimoto (1997). "Clostridium Perfringens Enterotoxin Utilizes Two Structurally Related Membrane Proteins as Functional Receptors in Vivo." J Biological Chemistry **272**: 26652-26658.

Klyce, S. D. and R. W. Beuerman (1988). "Structure and Function of the Cornea." The Cornea: 3-23.

Lambert, L. A., W. A. Chambers, S. Green, K. C. Gupta, R. N. Hill, P. M. Hurley, et al. (1993). "The Use of Low-volume Dosing in the Eye Irritation Test." Food and Chemical Toxicology **31**: 99-103.

Liebsch, M., D. Traue, C. Barrabas, H. Spielmann, P. Uphill, S. Wilkins, et al. (2000). "The ECVAM Prevalidation Study on the Use of EpiDerm for Skin Corrosivity Testing." ATLA **28**: 371-401.

Mann, I. and B. D. Pullinger (1942). "A Study of Mustard Gas Lesions of the Eyes of Rabbits and Men." Proc Roy Soc Med **35**: 229-244.

Marzulli, F. N. and D. I. Ruggles (1973). J Assoc Off Anal Chem **56**: 905.

Maurer, J. K., H. F. Li, W. M. Petroll, R. D. Parker, H. D. Cavanagh and J. V. Jester (1997). "Confocal Microscopic Characterization of Initial Corneal Changes of Surfactant-induced Eye Irritation in the Rabbit." Toxicol Appl Pharmacol **143**(2): 291-300.

Maurer, J. K. and R. D. Parker (1996). "Light Microscope Comparision of Surfactant Induced Eye Irritation i Rabbits and Rats at Three Hours and Recovery/day 35." Toxicologic Pathology **24**: 403-411.

Maurer, J. K., R. D. Parker and J. V. Jester (2002). "Extent of Initial Corneal Injury as the Mechanistic Basis for Ocular Irritation: Key Findings and Recommendations for the Development of Alternative Assays." Regul Toxicol Pharmacol **36**(1): 106-117.

Maurer, J. L. and R. D. Parker (1996). "LIght Microscope Comparison of Surfactant-induced Eye Irritation in Rabbits and Rats at Three Hours and Recovery/day 35." Toxicologic Pathology **6**: 464-477.

McConnell, H. M., J. C. Owicki, J. W. Parce, D. L. Miller, G. T. Baxter, H. G. Wada, et al. (1992). "The cytosensor microphysiometer: biological applications of silicon technology." Science **257**(5078): 1906-12.

Millichamp, N. J. (1999). "Species Specificity: Factors Affecting the Interpretation of Species Differences in Toxic Responses of Ocular Tissues. ." Ophthalmic Toxicology, Second Edition, Target Organ Toxicology Series (Ed. GCY Chiou). Philadelphia: Taylor and Francis, 89-117.

Mishima, S., A. Gasset, S. D. Klyce and J. L. Baum (1966). "Determination of Tear Volume and Tear Flow." Invest Ophthalmol Vis Sci **5**: 264-276.

NIEHS (2006). "Current Status of In Vitro Methods for Identifying Ocular Corrosives and Severe Irritations: Bovine Corneal Opactiy and Permeability Test Background Review Document." NIH Publication No. **06-4512**.

Prinsen, M. K. (2006). "The Draize Eye Test and In Vitro Alternatives; A Left-handed Marriage?" Toxicol In Vitro **20**(1): 78-81.

Rees, W. M., J. E. Swanson, J. Burdick, D. S. Hilgers and J. Harbell (2001). "Evaluating Toxic Synergism in Hypochlorite-containing Solutions Using the Bovien Cornea Opacity and Permeability (BCOP) Assay." The Toxicologist **60**: 99.

Roggeband, R., M. York, M. Pericoi and W. Braun (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**: 727-734.

Sina, J. (1994). "Validation of the Bovine Corneal Opacity-Permeability Assay as a Predictor of Ocular Irritation Potential." In Vitro Toxicol **7**(3): 283-290.

Sina, J. F., D. M. Galer, R. G. Sussman, P. Gautheron, E. V. Sargent, B. Leong, et al. (1995). "A collaborative Evaluation of Seven Alternatives to the Draize Eye Irritation Test Using Pharmaceutical Intermediates." Fund. Appl. Toxic. **26**: 20-31.

Stern, M., M. Klausner, R. Alvarado, K. J. Renskers and M. S. Dickens (1998). "Evaluation of the EpiOcular(TM) Tissue Model as an Alternative to the Draize Eye Irritation Test." Toxicology In Vitro **12**: 455-461.

Swanson, J. E. and J. Harbell (2000). "Evaluating the Eye Irritancy Potential of Ethanolic Test Materials with the Bovine Corneal Opacity and Permeability Assay." The Toxicologist **54**(1): 188.

Swanson, J. E., L. K. Lake, T. A. Donnelly, J. Harbell and J. Huggins (1995). "Prediction of Ocular Irritation of Full-Strength Cleaners and Strippers by Tissue Equivalent and Bovine Corneal Assays." Toxicology - Cutaneous and Ocular Toxicology **14**(3): 179-195.

Swanson, J. E., B. T. White, B. P. Gran, J. Merrill and J. Harbell (2003). "Evaluating Oxidizing/reactive Cleaning Products in the Bovine Corneal Opacity and Permeability (BCOP) Assay." The Toxicologist **72**: 220-221.

Swanston, D. W. (1985). "Assessment of the Validity of Animal Techniques in Eye Irritation Testing." Food and Chemical Toxicology **23**: 169-173.

Thakur, A., A. Clegg and e. al (1997). "Modulation of Cytokine Production from an EpiOcular Corneal Cell Culture Model in Response to Staphylococcus Aureus Superantigen." Aust N Z J Ophthalmol **25 Suppl 1**: S43-45.

Ubels, J. L. (1998). "Effect of Hydration on Opacity in the Bovine Corneal Opacity and Permeability (BCOP) Assay." J. Toxicol. Cut & Ocular Toxicol. **17**(4): 197-220.

UN (2003). "Globally Harmonised System of Classifcation and Labelling of Chemicals (GHS)." New York & Geneva: United Nations Publications.

Van Meer, G., W. van Hof and I. nan Genderen (1992). "Tight Junctions and Polarity of Lipids " Tight Junctions (ed Cereijido M.): 187-201.

Walker, A. P. (1985). "A More Realistic Animal Technique for Predicting Human Eye Response." Food Chem Toxicol **23**: 175-178.

Weil, C. S. and R. A. Scala (1971). "Study of Intra- and Interlaboratory Variability in the Results of Rabbit Eye and Skin Irritation Tests." Toxicol Appl Pharmacol **19**(2): 276-360.

Wilhelmus, K. R. (2001). "The Draize Eye Test." Surveys in Ophthalmology **45**: 493-515.