

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

DRAFT PROPOSAL FOR AN UPDATE OF TEST GUIDELINE 431

***In Vitro* Skin Corrosion: Reconstructed Human Epidermis (RhE) Test Method**

INTRODUCTION

1. Skin corrosion refers to the production of irreversible damage to the skin manifested as visible necrosis through the epidermis and into the dermis, following the application of a test material [as defined by the United Nations (UN) Globally Harmonised System of Classification and Labelling of Chemicals (GHS)] (1). This Test Guideline provides an *in vitro* procedure allowing the identification of corrosive chemical substances and mixtures.

2. The assessment of skin corrosivity has typically involved the use of laboratory animals (OECD Test Guideline 404 (TG 404); adopted in 1981 and revised in 1992 and 2002)(2). In relation to animal welfare TG 404 was revised in 2002, allowing for the determination of skin corrosion by applying a tiered testing strategy, using validated *in vitro* or *ex vivo* test methods, thus avoiding pain and suffering of animals. In addition to TG 431 (originally adopted in 2004)(3), two other *in vitro* test methods for testing of corrosivity have been validated and adopted as OECD Test Guidelines 430 (4) and 435 (5).

3. This Test Guideline is based on a reconstructed human epidermis (RhE) model, which in its overall design (the use of human derived epidermis keratinocytes as cell source and use of representative tissue and cytoarchitecture) closely mimics the biochemical and physiological properties of the upper parts of the human skin, *i.e.*, the epidermis. This updated Test Guideline also includes a set of Performance Standards (PS)(Annex 1) for the assessment of similar and modified RhE-based test methods (6), in accordance with the principles of Guidance Document No. 34 (7).

4. Prevalidation studies (8), followed by a formal validation study of *in vitro* methods for assessing skin corrosion (9)(10) have been conducted (11)(12). The outcome of these studies and other published literature (13) led to the recommendation that the following tests could be used for regulatory purposes for the assessment of *in vivo* skin corrosivity (14)(15)(16): the transcutaneous electrical resistance test (see Test Guideline 430) and the human skin model test (this Guideline).

5. Before a proposed similar or modified *in vitro* RhE test method for skin corrosion can be used for regulatory purposes, its reliability, relevance (accuracy), and limitations for its proposed use should be determined to ensure that it is similar to that of the TG 431, in accordance with the requirements of the PS set out in this Test Guideline (Annex 1).

DEFINITIONS

6. Definitions used are provided in Annex 2.

43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90

INITIAL CONSIDERATIONS

7. Validation studies have reported that tests employing human skin models (8)(9)(10)(13) are able to discriminate between known skin corrosives and non-corrosives with an overall sensitivity of 82% (23/28) and specificity of 84% (27/32) for a database of 60 substances. The test protocol may also provide an indication of the distinction between severe and less severe skin corrosives.

8. The test described in this Test Guideline allows the identification of corrosive chemical substances and mixtures. It further enables the identification of non-corrosive substances and mixtures when supported by a weight of evidence determination using other existing information (e.g., pH, structure-activity relationships, human and/or animal data) (1)(2)(17)(18). It does not normally provide adequate information on skin irritation, nor does it allow the subcategorisation of corrosive substances as permitted in the Globally Harmonised Classification System (GHS) (1).

9. This Test Guideline also includes a set of PS (Annex 2) for determining the validation status of new and revised skin corrosion test methods that are structurally and mechanistically similar to validated RhE test methods (6), in accordance with the principles of Guidance Document No. 34 (7). These performance standards include a list of 24 reference chemicals by which to evaluate assay performance, the essential test method components that should be included in the protocol for the test method to be considered structurally and mechanistically similar, and the minimum accuracy and reliability necessary for the test method to be considered comparable to validated RhE test methods. Within the reference chemical list, a subset of 12 proficiency chemicals (Annex 2) is provided that can be used by laboratories to demonstrate proficiency in using *in vitro* human skin models.

10. For a full evaluation of local skin effects after single dermal exposure, it is recommended to follow the sequential testing strategy as appended to TG 404 (2) and provided in the Globally Harmonised System (1). This testing strategy includes the conduct of *in vitro* tests for skin corrosion (as described in this guideline) and skin irritation before considering testing in live animals.

PRINCIPLE OF THE TEST

11. The test material is applied topically to a three-dimensional RhE model, comprised of normal, human-derived epidermal keratinocytes, which have been cultured to form a multilayered, highly differentiated model of the human epidermis. It consists of organized basal, spinous and granular layers, and a multilayered *stratum corneum* containing intercellular lamellar lipid layers arranged in patterns representing main lipid classes analogous to those found *in vivo*.

12. The principles of the RhE test method is based on the premise that corrosive chemicals are able to penetrate the *stratum corneum* by diffusion or erosion, and are cytotoxic to the cells in the underlying layers. Cell viability is measured by dehydrogenase conversion of the vital dye MTT [3-(4,5-Dimethyl-2-thiazol-2-yl)-2,5-diphenyl-2H tetrazolium bromide, Thiazolyl blue; EINECS number 206-069-5, CAS number 298-93-1], into a blue formazan salt that is quantitatively measured after extraction from tissues (19). Corrosive chemicals are identified by their ability to decrease cell viability below defined threshold levels (*i.e.*, $\leq 50\%$).

DEMONSTRATION OF PROFICIENCY

13. Prior to routine use of any RhE test method that adheres to this Test Guideline, laboratories should demonstrate technical proficiency, using the twelve Proficiency Chemicals recommended in Table 1. For similar tests developed under this Test Guideline that are structurally and mechanistically similar to the RhE test method, the PS requirements described in Annex 1 of this Test Guideline should be used to demonstrate similar reliability and accuracy of the test method prior to its use using the test method for regulatory testing.

14. As part of the proficiency exercise, it is recommended that the user verify the barrier properties of the tissues after receipt as specified by the test method RhE model manufacturer. Once a test method has been successfully established and proficiency in its use has been demonstrated, such verification will not be necessary on a routine basis. However, when using a test method routinely, it is recommended to continue to assess the barrier properties in regular intervals, *e.g.*, every six or twelve months.

Table 1. Proficiency Chemicals

| Chemical | CASRN | UN In Vivo PG | pH ¹ |
|-------------------------------|------------|---------------|-----------------|
| 1,2-Diaminopropane | 78-90-0 | II | 8.3 |
| Dimethyldipropylenetriamine | 10563-29-8 | I | 8.3 |
| 2-tert-Butylphenol | 88-18-6 | II/III | 3.9 |
| Potassium hydroxide (10% aq.) | 1310-58-3 | II | 13.1 |
| Sulfuric acid (10%) | 7664-93-9 | II/III | 1.2 |
| Octanoic acid (caprylic acid) | 124-07-2 | II/III | 3.6 |
| 4-Amino-1,2,4-triazole | 584-13-4 | NC | 5.5 |
| Eugenol | 97-53-0 | NC | 3.7 |
| Phenethyl bromide | 103-63-9 | NC | 3.6 |
| Tetrachloroethylene | 127-18-4 | NC | 4.5 |
| Isostearic acid | 30399-84-9 | NC | 3.6 |
| 4-(Methylthio)benzaldehyde | 3446-89-7 | NC | 6.8 |

¹ The pH values were obtained from Fentem et al. (1998) and Barratt et al. (1998).

Most of the chemicals listed are taken from the list of chemicals selected for the ECVAM international validation study (9). Their selection is based on the following criteria:

- i) equal number of corrosive and non-corrosive substances;
- ii) commercially available substances covering most of the relevant chemical classes;
- iii) inclusion of severely corrosive as well as less corrosive substances in order to enable discrimination based on corrosive potency;
- iv) choice of chemicals that can be handled in a laboratory without posing other serious hazards than corrosivity.

PROCEDURE

Human skin models

15. Human skin models can be constructed or obtained commercially (*e.g.*, the EpiDerm™ and

126 EPISKIN™ models) (19)(20)(21)(22) or be developed or constructed in the testing laboratory (23)(24). It
127 is recognised that the use of human skin is subject to national and international ethical considerations and
128 conditions. Any new model should be validated (at least to the extent described in paragraph 11). Human
129 skin models used for this test should comply with the following:

130

131 **General model conditions:**

132

133 16. Normal human keratinocytes should be used to reconstruct the epithelium. Multiple layers of
134 viable epithelial cells (basal layer, *stratum spinosum*, *stratum granulosum*) should be present under a
135 functional *stratum corneum*. *Stratum corneum* should be multilayered containing the essential lipid profile
136 to produce a functional barrier with robustness to resist rapid penetration of cytotoxic marker chemicals,
137 e.g., sodium dodecyl sulphate (SDS) or Triton X-100. The barrier function should be demonstrated and
138 may be assessed either by determination of the concentration at which a marker chemical reduces the
139 viability of the tissues by 50% (IC₅₀) after a fixed exposure time, or by determination of the exposure time
140 required to reduce cell viability by 50% (ET₅₀) upon application of the marker chemical at a specified,
141 fixed concentration. The containment properties of the RhE model should prevent the passage of material
142 around the *stratum corneum* to the viable tissue, which would lead to poor modelling of skin exposure. The
143 RhE model should be free of contamination by bacteria, viruses, mycoplasma, or fungi.

144

145 **Functional model conditions:**

146

147 17. The magnitude of viability is usually quantified by using MTT or other metabolically converted
148 vital dyes. In these cases the optical density (OD) of the extracted (solubilized) dye from the negative
149 control tissue should be at least 20 fold greater than the OD of the extraction solvent alone [for an
150 overview, see (25)]. The negative control tissue should be stable in culture (provide similar viability
151 measurements) for the duration of the test exposure period. The stratum corneum should be sufficiently
152 robust to resist the rapid penetration of certain cytotoxic marker chemicals (e.g., 1% Triton X-100). This
153 property can be estimated by the exposure time required to reduce cell viability by 50% (ET₅₀) (e.g. for
154 the EpiDerm™ and EPISKIN™ models this is > 2 hours). The tissue should demonstrate reproductivity
155 over time and preferably between laboratories. Moreover it should be capable of predicting the corrosive
156 potential of the proficiency chemicals (Table 1) when used in the testing protocol selected.

157

158 **Application of the test and control substances**

159

160 18. Two tissue replicates are used for each treatment (exposure time), including controls. For liquid
161 materials, sufficient test substance should be applied to uniformly cover the skin surface; a minimum of
162 25µL/cm² should be used. For solid materials, sufficient test substance should be applied evenly to cover
163 the skin, and it should be moistened with deionised or distilled water to ensure good contact with the
164 skin. Where appropriate, solids should be ground to a powder before application. The application method
165 should be appropriate for the test substance (see e.g., reference 10). At the end of the exposure period, the
166 test material should be carefully washed from the skin surface with an appropriate buffer, or 0.9% NaCl.

167

168 19. Concurrent NC and positive controls (PC) should be used in each run to demonstrate that
169 viability (NC), barrier function and resulting tissue sensitivity (PC) of the tissues are within a defined
170 historical acceptance range. The suggested positive control substances are glacial acetic acid or 8N KOH.
171 The suggested negative controls are 0.9% NaCl or water.

172

173 **Cell viability measurements**

174

175 20. The assay used for determining the magnitude of viability is the MTT (19), if another test method

176 is used its validity would have to be shown prior to use under this Test Guideline. The optical density (OD)
177 of the extracted (solubilised) dye from the tissue treated with the negative control (NC) should be at least
178 20 fold greater than the OD of the extraction solvent alone. It should be documented that the tissue treated
179 with NC is stable in culture (provide similar viability measurements) for the duration of the test exposure
180 period.

181

182 21. Optical properties of the test chemical or its chemical action on the MTT may interfere with the
183 assay leading to a false estimate of viability (because the test chemical may prevent or reverse the colour
184 generation as well as cause it). This may occur when a specific test chemical is not completely removed
185 from the tissue by rinsing or when it penetrates the *epidermis*. If the test chemical acts directly on the
186 MTT, is naturally coloured, or becomes coloured during tissue treatment, additional controls should be
187 used to detect and correct for test chemical interference with the viability measurement technique. Detailed
188 description of how to correct direct MTT reduction and interferences by colouring agents is available in the
189 SOPs for the three validated test methods (22)(23)(24). Non specific colour (NSC) due to these
190 interferences should not exceed 30% of NC (for corrections). If NSC > 30%, the test chemical is
191 considered as incompatible with the test method.

192

193 **Interpretation of results**

194

195 22. The OD values obtained for each test sample can be used to calculate percent viability relative to
196 the negative control, which is arbitrarily set at 100%. The cut-off percentage cell viability value
197 distinguishing corrosive from non-corrosive test materials (or discriminating between different corrosive
198 classes), or the statistical procedure(s) used to evaluate the results and identify corrosive materials, should
199 be clearly defined and documented, and be shown to be appropriate. In general, these cut-off values are
200 established during test optimisation, tested during a prevalidation phase, and confirmed in a validation
201 study. As an example, the prediction of corrosivity associated with the EpiDerm™ model is (13):

202

203 23. The test substance is considered to be corrosive to skin:

204 i) if the viability after 3 minutes exposure is less than 50%, or

205

206 ii) if the viability after 3 minutes exposure is greater than or equal to 50 % and the viability after
207 1 hour exposure is less than 15%.

208

209 24. The test substance is considered to be non-corrosive to skin:

210

211 i) if the viability after 3 minutes exposure is greater than or equal to 50% and the viability after 1
212 hour exposure is greater than or equal to 15%.

213

214

215 **DATA AND REPORTING**

216

217 **Data**

218

219 25. For each test, data from individual replicate tissues (*e.g.*, OD values and calculated percentage
220 cell viability data for each test chemical, including classification) should be reported in tabular form,
221 including data from repeat experiments as appropriate. In addition means \pm SD for each test should be
222 reported. Observed interactions with MTT reagent and coloured test chemicals should be reported for each
223 tested chemical.

224

225 **Test report**

- 226
227 26. The test report should include the following information:
228 Test and Control Chemicals:
229 -Chemical name(s) such as CAS name and number, if known;
230 -Purity and composition of the chemical (in percentage(s) by weight);
231 -Physical-chemical properties relevant to the conduct of the study (*e.g.*, physical state,
232 stability, volatility and pH, water solubility if known);
233 -Treatment of the test/control chemicals prior to testing, if applicable (*e.g.*, warming,
234 grinding);
235 -Storage conditions;
- 236 Justification of the RhE model and protocol used:
237 Test Conditions:
238 - Cell system used;
239 - Calibration information for measuring device, and bandpass used for measuring cell viability
240 (*e.g.*, spectrophotometer);
241 - Complete supporting information for the specific RhE model used including its performance.
242 This should include, but is not limited to;
243 i) Viability
244 ii) Barrier function
245 iii) Morphology
246 iv) Reproducibility and predictivity
247 v) Quality controls (QC) of the model
248 - Details of the test procedure used;
249 - Test doses used, duration of exposure and post treatment incubation period;
250 - Description of any modifications of the test procedure;
251 - Reference to historical data of the model. This should include, but is not limited to;
252 i) acceptability of the QC data with reference to historical batch data
253 ii) acceptability of the positive and negative control values with reference to positive and
254 negative control means and ranges
255 - Description of evaluation criteria used including the justification for the selection of the cut-
256 off point(s) for the prediction model
- 257 Results:
258 - Tabulation of data from individual test substances;
259 - Description of other effects observed
- 260 Discussion of the results
261
262 Quality assurance statement for Good Laboratory Practice compliant studies:
263 – statement should indicate all inspections made during the study and the dates any
264 results were reported to the Study Director. The statement should also confirm that
265 the final report reflects the raw data.
266

| | |
|-----|------------|
| 267 | Conclusion |
| 268 | |
| 269 | |
| 270 | |
| 271 | |

272

273

LITERATURE

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

- (1) United Nations (UN) (2007). Globally Harmonized System of Classification and Labelling of Chemicals (GHS), Second revised edition, UN New York and Geneva, 2007. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev02/02files_e.html]
- (2) OECD (2004). Test Guideline 404. OECD Guideline for the Testing of Chemicals. Acute Dermal Irritation/Corrosion. Updated Guideline, adopted April 24, 2002. Available at: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html]
- (3) OECD (2004). Test Guideline 431. OECD Guideline for the Testing of Chemicals. In Vitro Skin Model. Adopted April 13, 2004. Available at: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html]
- (4) OECD (2004). Test Guideline 430. OECD Guideline for the Testing of Chemicals. In Vitro Skin Corrosion: Transcutaneous Electrical Resistance (TER). Adopted April 13, 2004. Available at: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html]
- (5) OECD (2006). Test Guideline 435. OECD Guideline for the Testing of Chemicals. In Vitro Membrane Barrier Test Method. Adopted July 19, 2006. Available at: [http://www.oecd.org/document/22/0,2340,en_2649_34377_1916054_1_1_1_1,00.html]
- (6) ICCVAM 2004. Recommended Performance Standards for *In Vitro* Test Methods for Skin Corrosion. NIH Publication Number 04-4510. Research Triangle Park, NC: National Institute of Environmental Health Sciences. Available at: http://iccvam.niehs.nih.gov/docs/dermal_docs/ps/ps044510.pdf
- (7) OECD (2005). OECD Series on Testing and Assessment No. 34. Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Available at: [http://www.oecd.org/document/30/0,3343,en_2649_34377_1916638_1_1_1_1,00.html]
- (8) Botham, P.A., Chamberlain, M., Barratt, M.D., Curren, R.D., Esdaile, D.J., Gardner, J.R., Gordon, V.C., Hildebrand, B., Lewis, R.W., Liebsch, M., Logemann, P., Osborne, R., Poniec, M., Regnier, J.F., Steiling, W., Walker, A.P., and Balls, M. (1995). A prevalidation study on *in vitro* skin corrosivity testing. The report and recommendations of ECVAM Workshop 6. *ATLA* 23, 219-255.
- (9) Barratt, M.D., Brantom, P.G., Fentem, J.H., Gerner, I., Walker, A.P., and Worth, A.P. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 1. Selection and distribution of the test chemicals. *Toxic. in Vitro* 12, 471-482.
- (10) Fentem, J.H., Archer, G.E.B., Balls, M., Botham, P.A., Curren, R.D., Earl, L.K., Esdaile, D.J., Holzhutter, H.-G., and Liebsch, M. (1998). The ECVAM international validation study on *in vitro* tests for skin corrosivity. 2. Results and evaluation by the Management Team. *Toxic. in Vitro* 12, 483-524.
- (11) Balls, M., Blaauboer, B.J., Fentem, J.H., Bruner, L., Combes, R.D., Ekwall, B., Fielder, R.J., Guillouzo, A., Lewis, R.W., Lovell, D.P., Reinhardt, C.A., Repetto, G., Sladowski, D., Spielmann, H., and Zucco, F. (1995). Practical aspects of the validation of toxicity test procedures. The report

- 322 and recommendations of ECVAM workshops. ATLA 23, 129-147.
323
- 324 (12) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (1997).
325 Validation and Regulatory Acceptance of Toxicological Test Methods. NIH Publication No. 97-
326 3981. National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA.
327 <http://iccvam.niehs.nih.gov/docs/guidelines/validate.pdf>.
328
- 329 (13) Liebsch, M., Traue, D., Barrabas, C., Spielmann, H., Uphill, P., Wilkins, S., McPherson, J.P.,
330 Wiemann, C., Kaufmann, T., Remmele, M. and Holzhütter, H.-G. (2000). The ECVAM
331 prevalidation study on the use of EpiDerm for skin corrosivity testing. ATLA 28, pp. 371-401.
332
- 333 (14) ECVAM (1998). ECVAM News & Views. ATLA 26, 275-280.
334
- 335 (15) ECVAM (2000). ECVAM News & Views. ATLA 28, 365-67.
336
- 337 (16) ICCVAM (Interagency Coordinating Committee on the Validation of Alternative Methods). (2002).
338 ICCVAM evaluation of EpiDermTM, EPISKINTM (EPI-200), and the Rat Skin Transcutaneous
339 Electrical Resistance (TER) assay: In Vitro test methods for assessing dermal corrosivity potential of
340 chemicals. NIH Publication No. 02-4502. National Toxicology Program Interagency Center for the
341 Evaluation of Alternative Toxicological Methods, National Institute of Environmental Health
342 Sciences, Research Triangle Park, NC, USA.
343 http://iccvam.niehs.nih.gov/methods/epiddocs/epis_brd.pdf
344
- 345 (17) OECD (2002) Extended Expert Consultation Meeting on The In Vitro Skin Corrosion Test
346 Guideline Proposal, Berlin, 1st –2nd November 2001, Secretariat's Final Summary Report, 27th
347 March 2002, OECD ENV/EHS, available upon request from the Secretariat.
348
- 349 (18) Worth AP, Fentem JH, Balls M, Botham PA, Curren RD, Earl LK, Esdaile DJ, Liebsch M (1998).
350 An Evaluation of the Proposed OECD Testing Strategy for Skin Corrosion. ATLA 26: 709-720.
351
- 352 (19) Tinois E, Gaetani Q, Gayraud B, Dupont D, Rougier A, Pouradier DX (1994). The Episkin model:
353 Successful reconstruction of human epidermis in vitro. In In vitro Skin Toxicology. Edited by A
354 Rougier, AM Goldberg and HI Maibach: 133-140.
- 355 (20) Cannon, C. L., Neal, P. J., Southee, J. A., Kubilus, J., and Klausner, M., 1994. New epidermal
356 model for dermal irritancy testing. Toxic. in Vitro 8, 889 - 891.
- 357 (21) Ponec, M., Boelsma, E., Weerheim, A., Mulder, A., Bouwstra, J., and Mommaas, M., 2000.
358 Lipid and ultrastructural characterization of reconstructed skin models. International Journal of
359 Pharmaceutics. 203, 211 - 225.
360
- 361 (22) Tinois E, Tiollier J, Gaucherand M, Dumas H, Tardy M, Thivolet J (1991). In vitro and post –
362 transplantation differentiation of human keratinocytes grown on the human type IV collagen film of
363 a bilayered dermal substitute. Experimental Cell Research 193: 310-319.
364
- 365 (23) Parenteau, N.L., Bilbo, P., Molte, C.J., Mason, V.S., and Rosenberg, H. (1992). The organotypic
366 culture of human skin keratinocytes and fibroblasts to achieve form and function. Cytotechnology 9,
367 163-171.
368

- 369 (24) Wilkins, L.M., Watson, S.R., Prosky, S.J., Meunier, S.F., Parenteau, N.L. (1994). Development of a
370 bilayered living skin construct for clinical applications. *Biotechnology and Bioengineering* 43/8,
371 747-756.
372
- 373 (25) Marshall, N.J., Goodwin, C.J., Holt, S.J. (1995). A critical assessment of the use of microculture
374 tetrazolium assays to measure cell growth and function. *Growth Regulation* 5, 69-84.
375
- 376 (26) Liebsch, M, Traue, D, Barrabas, C, Spielmann, H, Uphill, P, Wilkins, S, et al. 2000. The ECVAM
377 prevalidation study on the use of EpiDerm for skin corrosivity testing. *ATLA* 28:371-401.
378
379

380 ANNEX 1
381 PERFORMANCE STANDARDS FOR EVALUATION OF THE VALIDATION STATUS OF
382 PROPOSED NEW OR MODIFIED IN VITRO TEST METHODS THAT ARE STRUCTURALLY
383 AND MECHANISTICALLY SIMILAR TO THE *IN VITRO* HUMAN SKIN MODEL SYSTEMS
384 FOR SKIN CORROSION

385 **INTRODUCTION**

386 1. The purpose of Performance Standards (PS) is to communicate the basis by which new test
387 methods, both proprietary (*i.e.*, copyrighted, trademarked, registered) and non-proprietary can be
388 determined to have sufficient accuracy and reliability for specific testing purposes. These PS, based on
389 validated and accepted test methods, can be used to evaluate the reliability and accuracy of other analogous
390 test methods (colloquially referred to as “me-too” tests) that are based on similar scientific principles and
391 measure or predict the same biological or toxic effect (7).

392 2. Prior to adoption of modified test methods, *i.e.*, proposed potential improvements to an approved
393 test method, there should be an evaluation to determine the effect of the proposed changes on the test’s
394 performance and the extent to which such changes affect the information available for the other
395 components of the validation process. Depending on the number and nature of the proposed changes, the
396 generated data and supporting documentation for those changes, they should either be subjected to the
397 same validation process as described for a new test, or, if appropriate, to a limited assessment of reliability
398 and relevance using established PS (7).

399 3. Similar (me-too) or modified test methods proposed for use under this Test Guideline should be
400 evaluated to determine their reliability and accuracy using chemicals representing the full range of the TG
401 431 corrosivity scores.

402 4. These PS are based on the US-ICCVAM PS (6) for evaluating the validity of new or modified
403 RhE test methods. The PS consist of essential test method components, recommended reference
404 substances, and standards for accuracy and reliability that the proposed test method should meet or exceed.

405 **I) Essential test method components**

406 5. To ensure that a modified human skin cell culture model system is structurally and mechanistically
407 similar to EPISKIN™ and EpiDerm™ (EPI-200) and measures the same biological effect, the following
408 components should be included in the test method protocol:

- 409 1. The physical components of the test method including the human skin model construct
- 410 2. Application of test substance
- 411 3. Criteria for appropriate control substances
- 412 4. Measurement of cytotoxicity
- 413 5. Interpretation of results including criteria used to distinguish between corrosive and
414 noncorrosive test substances

415 If any of these criteria are not met, then these performance standards cannot be used for validation of the
416 new or modified test method.

417 **II) Minimum list of reference substances**

- 418 6. ICCVAM identified 24 minimum required reference substances (12 noncorrosives, 12 corrosives)
419 that are included in the skin corrosivity performance standards.
- 420 • The list of reference substances provides a representative distribution of the 60 chemicals used in
421 the ECVAM validation study of EPISKIN™ (9)(10).
 - 422 • The list of reference substances covers the range of corrosivity responses obtained for the *in vivo*
423 rabbit skin reference test method.
 - 424 • The 24 reference chemicals include 23 of the 24 chemicals used to validate EPIDERM™ (EPI-
425 200), a test method structurally and functionally similar to EPISKIN™ (26).
 - 426 • A subset of the 24 reference chemicals (12 total; 6 noncorrosives, 6 corrosives) serves as
427 proficiency chemicals for RhE test methods; the names of these chemicals are bolded.
 - 428

429 **Table 2. Recommended Chemicals for Validation of New or Modified *In Vitro* Human Skin Model**
 430 **Corrosivity Test Methods**
 431

| Chemical ¹ | CASRN | Chemical Class ² | UN <i>In Vivo</i> PG | pH ³ |
|--|------------|-----------------------------|----------------------|-----------------|
| <i>In Vivo</i> Corrosives | | | | |
| Phosphorus tribromide | 7789-60-8 | inorganic acid | II | 1.0 |
| Sulfuric acid (10%) | 7664-93-9 | inorganic acid | II/III | 1.2 |
| Boron trifluoride dihydrate | 13319-75-0 | inorganic acid | II | 1.5 |
| Glycol bromoacetate (85%) | 3785-34-0 | electrophile | II/III | 2.0 |
| Caprylic acid | 124-07-2 | organic acid | II/III | 3.6 |
| 2-tert-Butylphenol | 88-18-6 | phenol | II/III | 3.9 |
| Dimethyldipropylenetriamine | 10563-29-8 | organic base | I | 8.3 |
| Dimethylisopropylamine | 996-35-0 | organic base | II/III | 8.3 |
| 1,2-Diaminopropane | 78-90-0 | organic base | I | 8.3 |
| n-Heptylamine | 111-68-2 | organic base | II/III | 8.4 |
| 2-Mercaptoethanol, sodium salt (45% aq.) | 37482-11-4 | inorganic base | II/III | 12.0 |
| Potassium hydroxide (10% aq.) | 1310-58-3 | inorganic base | II | 13.1 |
| <i>In Vivo</i> Noncorrosives | | | | |
| Sulfamic acid | 5329-14-6 | inorganic acid | NC | 1.5 |
| Isostearic acid | 30399-84-9 | organic acid | NC | 3.6 |
| Phenethyl bromide | 103-63-9 | electrophile | NC | 3.6 |
| Eugenol | 97-53-0 | phenol | NC | 3.7 |
| 1,9-Decadiene | 1647-16-1 | neutral organic | NC | 3.9 |
| <i>o</i> -Methoxyphenol | 90-05-1 | phenol | NC | 3.9 |
| Sodium lauryl sulfate (20% aq.) | 151-21-3 | surfactant | NC | 3.9 |
| Tetrachloroethylene | 127-18-4 | neutral organic | NC | 4.5 |
| 4-Amino-1,2,4-triazole | 584-13-4 | organic base | NC | 5.5 |
| 4-(methylthio)-Benzaldehyde | 3446-89-7 | electrophile | NC | 6.8 |
| Sodium carbonate (50% aq.) | 7664-93-9 | inorganic base | NC | 11.7 |
| Dodecanoic acid (lauric acid) | 143-07-7 | organic acid | NC | ND |

432 Abbreviations: aq = aqueous; CASRN = Chemical Abstracts Service Registry Number; PG = Packing
 433 Group; NC = Noncorrosive; ND = not determined (unable to measure); UN = United Nations.
 434 Recommended proficiency chemicals are indicated in bold type.

435 ¹These chemicals, sorted first by corrosives versus noncorrosives and then by pH, were selected from
 436 among the 60 chemicals used by ECVAM to validate EPISKIN™ (9)(10). Unless otherwise indicated, the
 437 chemicals were tested at the purity level obtained when purchased from a commercial source (9). The goal
 438 of the selection process is to include, to the extent possible, chemicals that: are representative of the range
 439 of corrosivity responses (e.g., noncorrosives; weak to strong corrosives) that the validated reference test
 440 method is capable of measuring or predicting; are representative of the chemical classes used in the
 441 validation process; reflect the performance characteristics of the validated reference test method; have
 442 chemical structures that are well-defined; induce reproducible results in the validated reference test
 443 method; induce definitive results in the *in vivo* reference test; are commercially available; and are not
 444 associated with prohibitive disposal costs.

445 ²Chemical class assigned by Barratt et al. (1998).

446 ³The pH values were obtained from Fentem et al. (1998) and Barratt et al. (1998).

447

448 **III) Standards for accuracy and reliability**

449

450 4. When evaluated using the minimum list of recommended reference chemicals, the proposed test
451 method should have reliability and performance (i.e., sensitivity, specificity, false positive rates, and false
452 negative rates) characteristics that are comparable to the performance of the validated reference test
453 method (10)(16).

454 5. An assessment of inter-laboratory reproducibility is not essential if the test method is to be used in
455 one laboratory only.

456 6. In terms of cell viability measurements, the median CV should not exceed 35% for studies
457 conducted in different laboratories (10)(16). The median CV for replicate studies conducted in the same
458 laboratory should be less than the median CV for studies conducted in different laboratories.

459

460

ANNEX 2

DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with “concordance” to mean the proportion of correct outcomes of a test method.

Skin corrosion *in vivo*: is the production of irreversible damage of the skin; namely, visible necrosis through the epidermis and into the dermis, following the application of a test substance for up to four hours. Corrosive reactions are typified by ulcers, bleeding, bloody scabs, and, by the end of observation at 14 days, by discoloration due to blanching of the skin, complete areas of alopecia, and scars. Histopathology should be considered to evaluate questionable lesions.

Cell viability: parameter measuring total activity of a cell population (e.g., as ability of cellular mitochondrial dehydrogenases to reduce the vital dye MTT), which, depending on the endpoint measured and the test design used, correlates with the total number and/or vitality of the cells.

Performance standards (PS): Standards, based on a validated test method, that provide a basis for evaluating the comparability of a proposed test method that is mechanistically and functionally similar. Included are; (i) essential test method components; (ii) a minimum list of Reference Chemicals selected from among the chemicals used to demonstrate the acceptable performance of the validated test method; and (iii) the similar levels of accuracy and reliability, based on what was obtained for the validated test method, that the proposed test method should demonstrate when evaluated using the minimum list of Reference Chemicals.

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method.

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility.

Sensitivity: The proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method.

Specificity: The proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method.

Tiered testing strategy: Testing which uses test methods in a sequential manner; the test methods selected in each succeeding level are determined by the results in the previous level of testing.