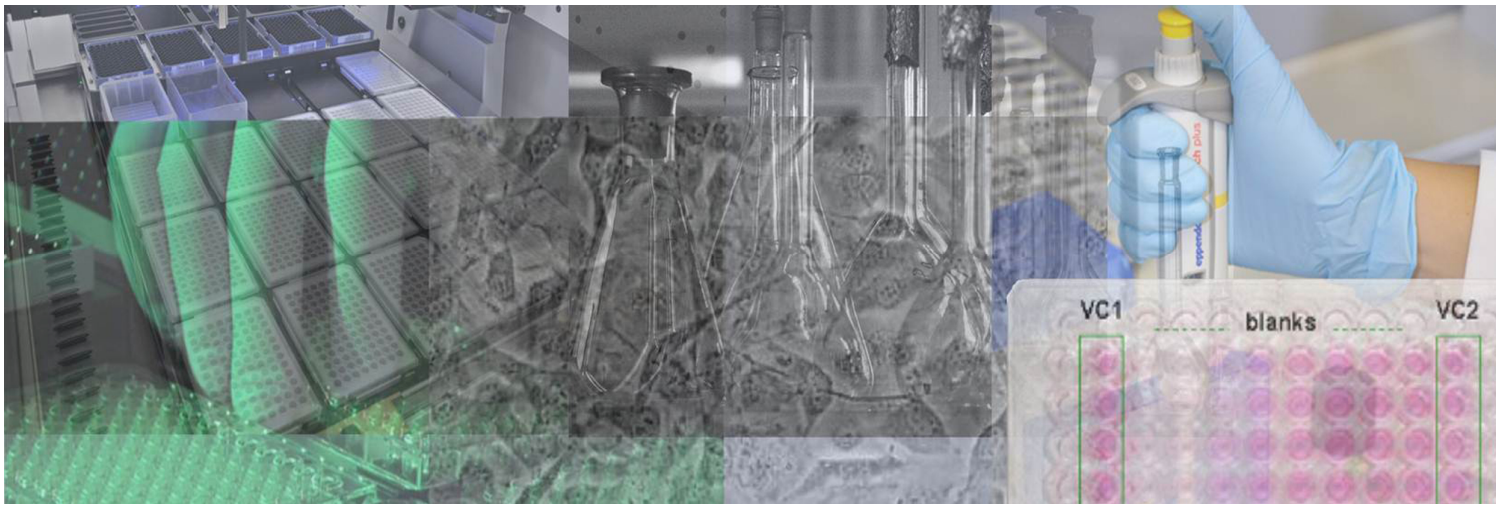
European Commission

JRC SCIENTIFIC AND POLICY REPORTS

EURL ECVAM Recommendation on the 3T3 Neutral Red Uptake Cytotoxicity Assay for Acute Oral Toxicity Testing

April 2013



Report EUR 25946 EN

Joint Research Centre

European Commission
Joint Research Centre
Institute for Health and Consumer Protection (IHCP)

Contact information
European Reference Laboratory for Alternatives to Animal Testing
Address: Joint Research Centre, Via Enrico Fermi 2749, TP xxx, 21027 Ispra (VA), Italy
E-mail: JRC-ECVAM-CONTACT@ec.europa.eu

http://ihcp.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

This publication is a Reference Report by the Joint Research Centre of the European Commission.

JRC79556

EUR 25946 EN

ISBN 978-92-79-29659-8 (pdf)
ISBN 978-92-79-29660-4 (print)

ISSN 1018-5593 (print)
ISSN 1831-9424 (online)

doi:10.2788/88799

Luxembourg: Publications Office of the European Union, 2013

Printed in Ispra

EUROPEAN COMMISSION
JOINT RESEARCH CENTRE

Institute for Health and Consumer Protection
European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)

EURL ECVAM RECOMMENDATION

on the 3T3 Neutral Red Uptake (3T3 NRU) Cytotoxicity Assay for the Identification of Substances not Requiring Classification for Acute Oral Toxicity

April 2013

ACKNOWLEDGEMENTS

This Recommendation was prepared by EURL ECVAM, part of the Institute for Health and Consumer Protection, Directorate-General Joint Research Centre of the European Commission.

The main contributors were Pilar Prieto (validation study coordinator), Claudius Griesinger (EURL ECVAM coordinator for ESAC peer reviews & ECVAM Recommendations), Patric Amcoff and Maurice Whelan. Valuable input was also received from Elisabet Berggren, Valérie Zuang, Susanne Belz and Gerry Bowe.

This Recommendation benefitted from comments and suggestions received from members of PARERE (EURL ECVAM's advisory body for Preliminary Assessment of Regulatory Relevance that brings together representatives of Member State regulatory bodies as well as EU agencies including ECHA, EFSA and EMA), and ESTAF (EURL ECVAM's Stakeholder Forum). Input was also provided by partner organisations of ICATM – the International Collaboration on Alternative Test Methods, and by the general public.

Comments or enquiries related to this EURL ECVAM Recommendation should be sent to:

European Union Reference Laboratory for Alternatives to Animal Testing
European Commission DG Joint Research Centre,
Institute for Health and Consumer Protection,
Via E. Fermi 2749,
I-21027 Ispra (VA), Italy.
e-mail: JRC-ECVAM-CONTACT@ec.europa.eu

TABLE OF CONTENTS

BACKGROUND TO EURL ECVAM RECOMMENDATIONS

The aim of a EURL ECVAM Recommendation is to provide EURL ECVAM views on the validity of the test method in question, to advice on possible regulatory applicability, limitations and proper scientific use of the test methods, and to suggest possible follow-up activities in view of addressing knowledge gaps.

During the development of its Recommendation EURL ECVAM consults with its advisory body for Preliminary Assessment of Regulatory Relevance (PARERE) and its EURL ECVAM Stakeholder Forum (ESTAF). Moreover, EURL ECVAM consults with other Commission services and its international validation partner organisations of the International Collaboration on Alternative Test Methods (ICATM). Moreover, before finalising its recommendations, EURL ECVAM invites comments from the general public and, if applicable, from the test method submitter.

EXECUTIVE SUMMARY

EURL ECVAM fully endorses the ESAC Opinion (Annex I) on the EURL ECVAM validation study to assess the predictive capacity of the 3T3 NRU in vitro cytotoxicity test method to identify substances not requiring classification for acute oral toxicity, based on a cut-off of LD_{50} > 2000 mg/kg b.w that corresponds to the upper threshold of UN GHS acute toxicity hazard Category 4 (United Nations, 2011). On the basis of the ESAC Opinion, EURL ECVAM makes the following recommendations:

(1) Considering the results of the EURL ECVAM study and the data available from the previous NICEATM/EURL ECVAM validation study (NIH, 2006), the 3T3 NRU test method shows a high sensitivity (ca 95%) and, consequently, a low false negative rate (ca 5%) when employed in conjunction with a prediction model to distinguish potentially toxic versus non-toxic (i.e. classified versus non-classified) substances. Therefore, substances found to be negative in this test would most likely not require classification for acute oral toxicity based on a cut-off value of >2000 mg / kg b.w.. Concluding on likely non-classification however requires careful consideration of the limitations of the assay (see points 2 and 3).

(2) The 3T3 NRU test method is sensitive to hazardous chemicals acting through general mechanisms of toxicity common to most cell types, often referred to as 'basal cytotoxicity'. However, chemicals not exhibiting significant cytotoxicity but which act through mechanisms specific only to certain cell types and tissues (e.g. of the heart or central nervous system) may not be indicated as potentially acutely toxic by this method. Moreover, chemicals requiring metabolic activation to induce toxicity may also go undetected since the cell model lacks significant metabolic capacity. Care must be taken therefore in interpreting negative results derived from this assay, despite the low false negative rate demonstrated in the EURL ECVAM validation study.

(3) Considering its limitations, results derived from the 3T3 NRU test method should be always used in combination with other information sources to build confidence in the decision not to classify a substance for acute oral toxicity. Possible complementary information sources include chemical analogues, physico-chemical properties, structural alerts, structure–activity relationships, and toxicokinetic data. The 3T3 NRU method would fit well therefore within a Weight of Evidence (WoE) approach or as a component of an Integrated Testing Strategy (ITS).

(4) The 3T3 NRU test method appears to be particularly relevant for the assessment of industrial chemicals since they are not designed to act on specific biological targets and in general tend not to be acutely toxic. Moreover, industrial chemicals which do exhibit toxicity are likely to act through multiple non-specific mechanisms that affect most cell types (i.e. basal cytotoxicity). Following the provisions of the REACH Regulation (1907/2006/EU) and in particular its Annex XI, data from the 3T3 NRU test method could be used within a WoE approach to adapt the standard information requirements.

(5) The 3T3 NRU test method has a high false positive rate therefore positive results cannot be readily used in a meaningful way. A likely reason is that the test method does not capture important biokinetic processes such as absorption, distribution, metabolism and excretion. Thus, certain chemicals, despite having cytotoxic potential, may not actually be acutely toxic via the oral route.

(6) As this test method informs about basal cytotoxicity which is a key event in many prevalent toxicological modes-of-action associated with both acute and chronic health effects, and since the EURL ECVAM validation study has shown the method to be amenable to automation and High Throughput Screening (HTS), it may constitute a valuable and economical information source for hazard profiling of substances.

(7) Respecting the provisions of Directive 2010/63/EU on the protection of animals used for scientific purposes, before embarking on animal experiments to identify acute oral toxicity the 3T3 NRU test method should be considered as an initial screening tool together with complimentary information in order to possibly reduce or avoid animal testing.

In summary, the 3T3 NRU test method could prove a valuable component of a WoE or ITS approach for supporting hazard identification and safety assessment in agreement with the EU CLP Regulation and international regulatory schemes implementing the upper threshold of UN GHS Category 4 as the cut-off for non-classification of substances. In particular, data from the 3T3 NRU assay may constitute an information source within a WoE approach under the provisions of the REACH regulation (Annex XI, 1.2) potentially supporting conclusions on absence of acute oral toxicity of industrial chemicals.

1. Introduction

1) The major regulatory driver for conducting acute oral toxicity studies is for classification and labelling (C&L) purposes to allow substances to be categorised according to their potential hazards and the dose required to cause toxicity (Creton et al., 2010; Seidle et al., 2010). In the EU, according to the provisions of Regulation 1272/2008 on Classification, Labelling and Packaging (the so-called CLP Regulation), substances with approximate LD_{50} values (or Acute Toxicity Estimates, ATE) of up to 2000 mg/kg b.w. are classified (according to Categories 1 to 4 of UN GHS). Substances with LD_{50} or ATE greater than 2000 mg/kg b.w. do not require classification as acute oral toxicants ('non-classified' / 'no Category').

2) Currently, there are only in vivo standardised test methods available to generate data on acute oral toxicity. In the frameworks of the Organisation for Economic Cooperation and Development (OECD) and the EU Test Methods Regulation (EC/440/2008), there are currently three internationally accepted guidelines: the Fixed Dose Procedure – FDP, TG420 (OECD, 2001a; EU test method B.1 bis), the Acute Toxic Class Method – ATC, TG423 (OECD, 2001b; EU test method B.1 tris), and the Up-and-Down Procedure – UDP, TG425 (OECD, 2001c). All these methods are sequential tests, where the outcome of the previous step/dose determines the next dose to be tested.

3) The possibility to use in vitro methods to predict acute oral toxicity has been investigated in several international projects, notably the Multicenter Evaluation of In Vitro Cytotoxicity (MEIC) programme (Ekwall et al., 1998; Ekwall et al., 2000), the Registry of Cytotoxicity (RC, Halle, 2003), and the NICEATM/EURL ECVAM In Vitro Basal Cytotoxicity International Validation Study (NIH, 2006), the latter employing also the 3T3 Neutral Red Uptake (NRU) assay. The MEIC results demonstrated that the majority of acutely toxic chemicals evaluated within the project appeared to exert their toxic effects through mechanisms of basal cytotoxicity, that is, by interfering with general cell functions common to all cells (Ekwall, 1999). In addition, most of these studies have shown a correlation of around 60 to 70% between in vitro IC_{50} cytotoxicity data and oral rat LD_{50} values, with higher correlation apparent in the lower toxicity range.

4) However, despite this good correlation, the NICEATM/EURL ECVAM validation study found that the neutral red uptake (NRU) basal cytotoxicity tests were not able to predict with sufficient accuracy the acute oral toxicity categories as defined by the United Nations Globally Harmonised System (UN GSH, 2011). The peer-review of this study concluded that the NRU basal cytotoxicity test methods, including the test performed with the 3T3 mouse embryonic fibroblast cell line (3T3 NRU), may be useful in a weight-of-evidence approach to determine the starting dose for acute oral in vivo toxicity tests (NIH, 2006, OECD, 2010).

5) As a follow-up of the previous NICEATM/EURL ECVAM validation study (see above), EURL ECVAM coordinated a second validation study in order to assess the capacity of the 3T3 NRU cytotoxicity assay to identify specifically non-classified chemicals on the basis of the 2000 mg/kg b.w. threshold introduced by the CLP Regulation. To this end, the study assessed the predictive capacity of the test method in conjunction with a dichotomous prediction model yielding only two categorical predictions instead of the six assessed in the previous NICEATM/EURL ECVAM study (five acute toxicity classes plus one test-outcome category relating to non-classified substances). These two predictions were I) potential 'negative' (i.e. predicted LD50 > 2000 mg/kg b.w.) and II) potential 'positive' (i.e. predicted LD50 • 2000 mg/kg b.w.. This was done on the basis of the hypothesis that the use of this test method, if sufficiently sensitive (i.e. low false negative rate), could significantly reduce in vivo testing for acute oral toxicity when used as a building block (i.e. as one of several data sources) within a tiered testing approach or an ITS to screen for potential negatives (i.e. substances not requiring classification) amongst general industrial chemicals. This rationale was further

supported by the apparent high prevalence of negatives (substances not requiring classification) amongst general industrial chemicals (Bulgheroni et al., 2009).

6) The EURL ECVAM follow-up study used the test method protocol and applied the millimole and weight IC_{50}-LD_{50} regressions, investigated in the previous NICEATM/EURL ECVAM validation study (NIH, 2006; Stokes et al., 2008), as well as a Receiver-Operator Characteristic (ROC) curve-based prediction model developed specifically for the follow-up study.

7) The follow-up study did not assess transferability or within/between laboratory reproducibility of the 3T3 NRU test method protocol, since these have been already validated during the previous NICEATM/EURL ECVAM validation study (NIH, 2006).

8) A secondary aim of the validation study was to assess whether two variants of the validated protocol would generate similar data on the basis of the test chemicals selected and to which extent they may be used for the purpose of identifying non-classified chemicals. The two protocol modifications were:
   a) a version of the 3T3 NRU protocol adapted to an automated High Throughput Screening (HTS) platform,
   b) an abbreviated version of the validated protocol that was targeted at evaluating acute oral toxicity levels around the 2000 mg/kg b.w. cut-off value.

9) After completion of the study and finalisation of the Validation Study Report (EC-ECVAM, 2011), EURL ECVAM requested the EURL ECVAM Scientific Advisory Committee (ESAC) at its meeting on 22 March 2011 to provide an ESAC Opinion on the study. An ESAC Working Group (WG) was established to review the results compiled in the Validation Study Report. Both the ESAC Opinion (EURL ECVAM, 2012a) (see Annex 1) and the ESAC WG report (EURL-ECVAM, 2012b) were adopted by ESAC on 23 March 2012.

Based on the above, EURL ECVAM developed the present Recommendation.

2. Test method definition

10) The 3T3 NRU cytotoxicity assay employs the BALB/c3T3 mouse fibroblast cell line and is based on the ability of viable cells to incorporate and bind the dye Neutral Red (NR). The uptake of NR is measured spectrophotometrically (Stokes et al., 2008). The basic concept of basal cytotoxicity assays is that chemicals exert their toxic effects by disrupting structures and functions universal to all cells, such as cell membranes (Gennari et al. 2004). With the basal cytotoxicity assays it is possible to quantify the cytotoxicity of a compound by its IC_{50} value, i.e. the concentration of the tested substance that decreases cell viability by 50% in the cell culture.

11) The 3T3 NRU assay addresses specifically the toxicity mechanism of basal cytotoxicity, since the fibroblast cell line does not capture mechanisms of action relating to interaction with specific molecular targets in certain tissues e.g. neuronal or cardiac receptors, ion channels etc. Furthermore, the cell line lacks metabolic competence associated with Phase I and Phase II biotransformation and therefore is sensitive to cytotoxicity induced by the parent compound and not of its metabolites.

12) The complete protocol of the 3T3 NRU test method is publicly available via EURL ECVAM's DataBase for ALternative Methods (DB-ALM) as "DB-ALM Protocol n°139" at the address: http://ecvam-dbalm.jrc.ec.europa.eu. The protocol provides a comprehensive description of the method together

with all the necessary technical details needed by an end-user laboratory to implement it in a self-sufficient manner.

3. Overall performance of the 3T3 NRU

13) EURL ECVAM concludes that the 3T3 NRU method has a high sensitivity (between 92-96%) and hence a low false negative rate (between 4-8%) for all three protocols used in the validation study. In this context, the millimole regression used in the NICEATM/EURL ECVAM validation study proved to be the most appropriate prediction model yielding a sensitivity of about 95% and hence a very low false negative rate of about 5%. Predictions obtained with the weight regression and ROC prediction models were less accurate.

14) While the method does correctly predict the majority of positives, it also over-predicts many negatives as positives (i.e. it has a high false positive rate or low specificity). This limits the usefulness of positive test results and leads to a comparably low rate of true negatives identified as such (40-44%). In addition to the issue of poor specificity, positive results from the 3T3 NRU do not resolve the acute toxicity categories required by international classification and labelling schemes (UN GHS, EU CLP).

15) Most importantly however, negative test results from the 3T3 NRU are largely accurate, i.e. substances identified as negatives are indeed to a very high extent actual negatives (low false negative rate). Thus, using the regression-based prediction model, the assay should be considered for use within an integrated approach to contribute to the identification of substances that are likely not to require classification under the CLP Regulation.

16) As industrial chemicals used for a wide variety of purposes are normally not designed to act on specific biological targets, they are generally expected to be of rather low toxicity. A retrospective analysis of registered industrial chemicals conducted by EURL ECVAM confirmed this: 87% of the 4219 substances notified in the EU New Chemicals Database (notifications up to 2008) are not classified as acute oral toxicants (Bulgheroni et al. 2009). Combining this information with the low false negative rate of the 3T3 NRU assay, it is conceivable that the method may support the reliable identification (as negatives) of a considerable fraction of substances that currently require testing for acute oral toxicity. If the 3T3 NRU assay was used in conjunction with other information sources (e.g. QSARs or structural alerts informing about the likely mode of action, biokinetic modelling etc), the assay might contribute to a reduction of animal testing for acute oral toxicity by up to 40%. Notably this figure represents a best-case scenario based on the assumptions of i) a prevalence of 87% of substances being assessed not requiring classification, ii) the predictive performance (i.e. ca 5% false negative rate) as determined in the EURL ECVAM follow-up study (which may vary in practice depending on the nature of chemicals tested) and iii) the availability of other information sources addressing the limitations of the 3T3 NRU assay and hence supporting the use of negative test results for subsequent decision making within the strategy. Therefore, this figure may be considerably lower for specific classes of chemicals, depending on the actual prevalence of acutely toxic substances and the likely mode-of-action in a particular chemical inventory or class. However, it should be noted that in the area of acute oral toxicity testing there are currently no non-animal methods employed for generating data on the presence or absence of potential hazards. Thus, irrespective of the prevalence of negatives amongst chemicals, any test method allowing identification of even one acute toxicity category of substances (here negatives) will likely contribute to a reduction of animal testing.

4. Limitations

17) The results obtained in the follow-up EURL ECVAM validation study support the findings previously reported (NIH, 2006, OECD 2010) which restrict the applicability of the 3T3 NRU test method to soluble compounds. Chemicals with metabolic (bioactivation and detoxification) pathways, and chemicals which do not exhibit basal cytotoxicity but which may act through organ-specific physiological mechanisms-of-action (i.e. cardiac channels, specific hepatic proteins, etc) not present in the 3T3 test system, may be misclassified.

5. Suggested regulatory use

18) Based on the outcome of the EURL ECVAM validation study, the 3T3 NRU can be used in an integrated approach to support identification of non-classified substances when using a threshold of > 2000 mg/kg b.w.. This threshold is used for non-classification by EU CLP and other classification and labelling schemes (e.g. US Department of Labor's US Occupational Safety and Health Administration) that implement UN GHS acute toxicity (oral route) categories 1 to 4[1]. Examples of integrated approaches are WoE analysis combining a variety of available information in a non-prescriptive manner or ITS that provide more specific guidance on the use of individual information sources in a retro- and/or prospective manner to reach decisions. Notably, following the provisions of the REACH Regulation and its Annex XI, data from the 3T3 NRU test method could already be used within a WoE approach to adapt the standard information requirements.

19) Since the 3T3 NRU test method is incapable of modelling organ-specific mechanisms of action, the use of the method for hazard and risk assessment of substances should be restricted to chemicals that have not been designed or formulated to interact with molecular targets specific to certain cell types. Care must be taken in interpreting the results of this assay when there is information that indicates that the chemical needs metabolic activation to exert its toxicity or that the chemical is detoxified. In some instances, the information relating to the biokinetic profile of the chemical and its likely mechanism(s) of action may allow testing of any potentially active metabolite(s).

20) When data derived from this test method are used for the hazard assessment of a chemical, for example through a WoE approach, a read-across from chemical analogues or within an ITS, these data should be supplemented where possible with information relating to the biokinetic profile of the chemical and its likely mechanism(s) of action.

21) The application of the 2000 mg/kg b.w. threshold to the data from the previous NICEATM/EURL ECVAM validation study (72 chemicals of which 42% were pharmaceuticals and 22% pesticides) showed only one false negative result (digoxin, a plant alkaloid with a highly specific mechanism of action). Moreover, the application of this threshold to 187 BASF in-house substances (70% of them were agrochemical active ingredients or formulations with specific modes of action) published by Schrage et al. (2011) confirmed the high sensitivity of the test method and its low false negative rate (only two false negative results were obtained). Taken together the outcomes of these studies and the follow-up study, EURL ECVAM concludes that it is not sufficiently clear whether the applicability

[1] It should be noted that UN GHS provides the possibility to use one more additional category (Category 5) intended to enable the classification and labelling of substances of very low toxicity that "under certain circumstances may present a danger to vulnerable populations" (UN GHS, 2011). Category 5 is subject to specific and limiting criteria, such as, for example, that "reliable evidence is already available that indicates the LD50 (or LC50) to be in the range of Category 5" as well as limitations with regard to the generation of prospective testing data, e.g. animal testing to determine (suspected) Category 5 ranges is discouraged by UN GHS.

of the 3T3 NRU test method may be extended beyond substances acting only via cytotoxicity. On the other hand, the specificity of the test method with the particular inventory used by Schrage et al. (2011) was even lower than that observed during the EURL ECVAM follow-up study, resulting in a high false positive rate limiting the utility of the test method for screening out negatives.

6. Follow-up activities recommended by EURL ECVAM

(1) Weight of Evidence (WoE) approach: Although data from the 3T3 NRU assay could be used within a WoE approach for adapting the standard information requirements for acute oral toxicity as defined by REACH, effort needs to be invested in showing how this might be realised in practice. This should include the identification of relevant and complementary information sources (including 3T3 NRU data), the design of algorithms for determining relative weighting and arriving at a conclusion, and the elaboration of case studies to evaluate and demonstrate WoE approaches suitable for different regulatory contexts. Ideally, the development of WoE knowhow should be formulated as guidance to prospective end-users.

(2) Integrated Testing Strategies (ITS): In order to make best use of the information provided by the 3T3 NRU assay, EURL ECVAM recommends that Integrated Testing Strategies be developed aiming at full or at least partial identification of acute oral toxicity hazard according to the GHS (Categories 1 to 4). In particular, the main purpose of the ITS should be to address information requirements under REACH and CLP related to this endpoint. In addition to including the 3T3 NRU method as a component in various ITS, consideration should also be given to other in vitro methods addressing specific mechanisms of action relevant to acute systemic toxicity, and to computational methods such as (Q)SAR and Physiologically Based BioKinetic (PBBK) modelling.

(3) Adverse Outcome Pathways (AOP): Having more explicit knowledge available on the numerous modes-of-action that lead to acute systemic toxicity would be very valuable in the design and validation of integrated prediction methods. EURL ECVAM recommends therefore that effort be invested in gathering and organising mode-of-action knowledge related to this endpoint, for example through the use of the OECD's Adverse Outcome Pathway framework and guidance. Developing AOPs and identifying the key biological events that they comprise will provide a theoretical understanding of acute systemic toxicity useful to developers of test methods and other predictive tools, as well as to validation and regulatory bodies.

(4) The 3T3 NRU assay as an initial step to reduce animal testing: In the absence of accepted ITS or WoE approaches for acute oral toxicity testing and in agreement with the provisions of Directive 2010/63/EU on the protection of animals used for scientific purposes, data from the 3T3 NRU assay should always be considered together with necessary supplementary information as an initial testing step prior to conducting animal experiments for the purposes of hazard classification with respect to acute oral toxicity. Moreover, if animal experiments are to be conducted, results of the 3T3 NRU assay should be considered when determining the starting dose for Guideline studies, as described in the OECD Guidance Document 129. In such cases, the in vitro data should be submitted to the regulatory authority together with the in vivo data. In addition, EURL ECVAM would like to receive this data (in confidence if necessary) to add to its database on the 3T3 NRU method to continue its post-validation performance characterisation.

(5) Further performance assessment: applicability, limitations and predictive capacity: To support the development of an ITS applicable to a wide range of substances, characterisation of the test method's applicability domain, possible limitations and predictive capacity should continue. Preferably, existing data should be used in a retrospective analysis for this characterisation. Data

for compounds classified as being highly toxic in terms of acute oral toxicity (i.e. category 1 and category 2) as well as substances known to act on specific molecular targets (i.e. pharmaceuticals, pesticides, biocides etc.) are expected to be particularly useful for this purpose. Furthermore, attention should be given to address the assay's limitations, for example to select complementary assays that address organ-specific effects (Zurich et al. 2012) and metabolically activation.

(6) Protocol variants to be used for further characterisation of the assay's overall performance: If additional testing is conducted to expand the 3T3 NRU's applicability domain and the information on its predictive performance, it should be based either on the simplified or the automated protocol variant (Bouhifd et al., 2012). As reproducibility of the assay (including the simplified and automated protocols) has been satisfactorily demonstrated, it would seem sufficient to conduct any additional testing in single laboratories (i.e. without the need for further ring trials).

(7) Use of biokinetic data as a complementary source of information: While generating more in vitro cytotoxicity data with the 3T3 NRU test method, EURL ECVAM suggests complementing them with biokinetic data to help with the interpretation of negative assay results in particular and, where possible, also positive results.

(8) Solubility of test materials: The reliability of the 3T3 NRU test method is only assured when the chemical to be tested has been correctly solubilised. Users should make every effort to follow the solubility protocol (available from the NICEATM/ICCVAM website and described in Stokes et. al., 2008) and also to report any solubility issues together with the test result when used for regulatory purposes.

REFERENCES

Bouhifd, M., Bories, G., Casado, J., Coecke, S., Norlén, H., Parissis, N., Rodrigues, R.M., Whelan, M.P., 2012. Automation of an in vitro cytotoxicity assay used to estimate starting doses in acute oral systemic toxicity tests. Food Chem Toxicol. 50(6):2084-96.

Bulgheroni, A., Kinsner-Ovaskainen, A., Hoffmann, S., Hartung, T., Prieto, P., 2009. Estimation of acute oral toxicity using the No Adverse Effect Level (NOAEL) from the 28-day repeated dose toxicity studies in rats. Regulatory Toxicology and Regulatory Toxicology and Pharmacology 53, 16–19.

Creton, S., Dewhurst, I.C., Earl, L.K., Gehen, S.C., Guest, R.L., Hotchkiss, J.A., Indans, I., Woolhiser, M.R., Billington, R., 2010. Acute toxicity testing of chemicals – Opportunities to avoid redundant testing and use alternative approaches. Critical Reviews in Toxicology 40(1):50-83.

Ekwall, B., Barile, F.A., Castaño, A., Clemedson, C., Clothier, R.H., Dierickx, P., Ekwall, B., Ferro, M., Fiskesjö, G., Garza-Ocañas, L., Gómez-Lechón, M.J., Gülden, M., Hall, T., Isomaa, B., Kahru, A,, Kerszman, G., Kristen, U., Kunimoto, M., Kärenlampi, S., Lewan, L., Loukianov, A., Ohno, T., Persoone, G., Romert, L., Sawyer, T.W., Shrivastava, R., Segner, H., Stammati, A., Tanaka, N., Valentino, M., Walum, E., Zucco, F., 1998. MEIC Evaluation of Acute Systemic Toxicity. Part VI. The Prediction of Human Toxicity by Rodent LD50 Values and Results From 61 In Vitro Methods. Alt. Lab. Anim. 26: 617–658.

Ekwall, B., 1999. Overview of the final MEIC results: II. The in vitro-in vivo evaluation, including the selection of a practical battery of cell tests for prediction of acute lethal blood concentrations to humans. Toxicology In Vitro 13: 665-673.

Ekwall, B., Ekwall, B., Sjöström, M., 2000. MEIC evaluation of acute systemic toxicity: Part VIII. Multivariate partial least squares evaluation, including the selection of a battery cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. Alt. Lab. Anim. 28 (suppl 1):201-34.

EURL ECVAM (2012a) ESAC Opinion Based on the ESAC Peer Review of the ECVAM-coordinated follow-up study to assess the predictive capacity of the validated 3T3 Neutral Red Uptake cytotoxicity assay for acute oral toxicity testing.
Online (March 2013):http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations

EURL ECVAM (2012b) ESAC Working Group Peer Review Consensus Report on the ECVAM-coordinated follow-up study to assess the predictive capacity of the already validated Neutral Red Uptake cytotoxicity assay for acute oral toxicity testing.
Online (March 2013):http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations

European Commission, JRC, IHCP, EURL ECVAM, 2013. DataBase service on ALternative Methods to animal experimentation (DB-ALM), DB-ALM Protocol: BALB/C 3T3 Neutral Red Cytotoxicity Assay, Online (March 2013): http://ecvam-dbalm.jrc.ec.europa.eu

European Commission, JRC, IHCP, EURL ECVAM, 2011. Report of the Validation Management Team on the ECVAM Follow-up study on the predictive capacity of the 3T3 Neutral Red Uptake cytotoxicity assay to correctly identify substances not classified for acute oral toxicity under the EU CLP system (LD_{50} > 2 000 mg/kg). Online (March 2013): http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations

European Union, 2008. Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures, amending and repealing directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (REACH).

European Union 2008. Regulation (EC) No 440/2008 of 30 May 2008 laying down test methods pursuant to Regulation (EC) No 1907/2006 of the European Parliament and of the Council on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). Official Journal of the European Union L 142.

Gennari, A., van den Berghe, C., Casati, S., Castell, J., Clemedson, C., Coecke, S. et al. 2004. Strategies to replace in vivo avute systemic toxicity testing. The report and Recommendations of ECVAM Workshop 50. Alt. Lab. Anim. 32:437-459.

Halle, W., 2003. The Registry of Cytotoxicity: Toxicity Tesing in Cell Cultures to predict Acute Toxicity (LD50) and to reduce Tesing in Animals. Alt. Lab. Anim. 31: 89-198.

National Institutes of Health (NIH), 2006. Background Review Document (BRD): Validation of Neutral Red Uptake Test Methods NIH /In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity. Publication No. 07-4518, November 2006. Available from: http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm

Organisation for Economic Cooperation and Development (OECD), 2001a. Guideline for Testing of Chemicals, 420, Acute Oral Toxicity – Fixed Dose Method. Paris, France. (http://www.oecd.org)

Organisation for Economic Cooperation and Development (OECD), 2001b. Guideline for Testing of Chemicals, 423, Acute Oral Toxicity – Acute Toxic Class Method. Paris, France. (http://www.oecd.org)

Organisation for Economic Cooperation and Development (OECD), 2001c. Guideline for Testing of Chemicals, 425, Acute Oral Toxicity – Up-and-Down Procedure. Paris, France. (http://www.oecd.org)

Organisation for Economic Cooperation and Development (OECD), 2010. Guidance Document No 129 on using Cytotoxicity Tests to Estimate Starting Doses for Acute Oral Systemic Toxicity Tests. Paris, France. (http://www.oecd.org)

Schrage, A., Hempel, K., Schulz, M., Kolle, S.N., van Ravenzwaay, B., Landsiedel, R., 2011. Refinement and Reduction of Acute Oral Toxicity Testing: A Critical Review of the Use of Cytotoxicity Data. Alt. Lab. Anim. 39, 273-295

Seidle, T., Robinson, S., Holmes, T., Creton, S., Prieto, P., Scheel, J., Chlebus, M., 2010. Cross-sector review of drivers and available 3Rs approaches for acute systemic toxicity testing. Toxicological Sciences 116(2):382-396.

Stokes, W.S., Casati, S., Strickland, J., Paris, M., 2008. Neutral Red Uptake Cytotoxicity Tests for Estimating Starting Doses for Acute Oral Toxicity Tests. Current Protocols in Toxicology Unit 20.4, 1-20.

United Nations, 2011. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fourth revised edition. Part 3: "Health hazards". Available online: http://www.unece.org/trans/danger/publi/ghs/ghs_rev04/04files_e.html

Zurich, M.-G., Stanzel, S., Kopp-Schneider, A., Prieto, P., Honegger, P., 2012. Evaluation of aggregating brain cell cultures for the detection of acute organ-specific toxicity. Toxicol. in Vitro (in press), http://dx.doi.org/10.1016/j.tiv.2012.06.018.

Annex 1 ESAC Opinion

Opinion of the EURL ECVAM Scientific Advisory Committee (ESAC) based on the ESAC Peer Review of the ECVAM-coordinated follow-up study to assess the predictive capacity of the validated 3T3 Neutral Red Uptake cytotoxicity assay for acute oral toxicity testing

ESAC Opinion Nr.	2012-01
ESAC request Nr.	2011-02
Date of ESAC Opinion	20.3.2012

EXECUTIVE SUMMARY

It has long been appreciated that there is a significant correlation between acute systemic toxicity in vivo (expressed on the basis of population LD50 values) and the impairment of cell viability in vitro (referred to as in vitro cytotoxicity), which reflects the adverse effects of toxic substances on universal processes of cellular physiology and molecular biology, intrinsic to all eukaryotic cells. Despite this appreciable correlation however, such cytotoxicity assays were not found capable of replacing animal testing as they lack what might be called "predictive resolution", i.e. they are not, in reference to rodent LD50 data, able to predict with sufficient accuracy the different toxicity classes used to categorise substances of oral toxicity (=positives). This was convincingly demonstrated in a joint NICEATM/EURL ECVAM validation study which however also found that one of this cytotoxicity assays, the 3T3 NRU assay, can be used for determining the starting dose (NIH 2006; OECD 2010) for relevant in vivo experiments conducted to generate data satisfying regulatory classification requirements, despite the intrinsic high variability of LD50 values (Bass et al., 1982; Rowan, 1983) which question the meaningfulness of such high a number of classes for acute toxicity.

Continuing previous work on the potential usefulness of cytotoxicity measurements for the assessment of acute oral systemic toxicity (Seibert et al, 1996; Gennari et al., 2004), EURL ECVAM conducted a follow-up study to the previous NICEATM/EURL ECVAM study. The follow-up study focused specifically on the ability of the 3T3 NRU assay, within a tiered approach or Integrated Testing Strategy, to identify substances considered negative, i.e. substances with LD50 values >2000 mg per kg body weight (EU 2008). As, among notified industrial chemicals, a significant proportion (ca 87%) of substances has been found to be negative (Bulgheroni, 2009), tools capable of identifying negatives with sufficient accuracy would potentially allow reducing animal testing to a considerable extent, bringing also cost and time savings since cytotoxicity assays are comparably inexpensive and fast to conduct.

The 3T3 NRU assay is based on the BALB/c 3T3 mouse fibroblast cell line and uses the traditional dye toluylene red (=neutral red) as a vital stain and readout for cell viability: inhibition of the 3T3 cells' ability to take up and retain neutral red indicates impaired viability or, inversely, cytotoxicity (expressed as IC50 values, measured on the basis of the cellular population in vitro). Importantly, as the test system (fibroblast cell line) does not recapitulate modes of actions relating to specific molecular targets (e.g. neuronal or cardiac receptors, channels etc.), the mechanistic relevance of the 3T3 NRU assays appears, prima facie, limited to basal cytotoxicity. Due to this possible restriction and due to the previous finding that predicted LD50 values are not sufficiently accurate to allow classification of positives (NIH 2006), the test method had been suggested by EURL ECVAM for use in an Integrated Testing Strategy.

As the test method protocol has been found to yield reproducible results during the previous NICEATM/EURL ECVAM validation study, the follow-up study made use of this protocol as well as the validated prediction models. To assess to which extent the test method is amenable to rapid and cost-effective screening approaches two additional protocol variants were assessed: one version of the 3T3/NRU protocol adapted to an automated platform and an abbreviated version of the validated protocol targeted at resolving acute oral toxicities around the 2000 mg/kg cut-off value.

Following a request from EURL ECVAM to ESAC in March 2011 (EURL ECVAM 2011, Annex 2) for scientific advice and peer review of this study, the ESAC set up a Working Group (ESAC WG) charged with the detailed scientific peer review of this study.

After careful review of the study report (EURL ECVAM 2010) and considering the detailed ESAC WG peer review consensus report (EURL ECVAM 2012), the ESAC concludes, in agreement with ECVAM, that the 3T3 NRU assay has a sufficiently high sensitivity (about 95%) to be considered for use within an Integrated Testing Strategy (ITS) to identify substances likely not requiring classification. Since all three protocol variants assessed in the study showed comparable

predictive performance, the 3T3 NRU assays appears amenable to simplification and automation, making it potentially a candidate assay for cost-effective high-throughput screening approaches.

However, as the assay is based on a fibroblast cell line and does not model specific modes of action, the ESAC is of the opinion that suitable guidance needs to be developed to ensure the appropriate future use of the assay. Clearly, substances designed to act via a specific biological target mechanism of action should not be tested in the 3T3 NRU assay or eventual testing data on such substances should at least be treated with caution. Thus, the assay seems most suitable for assessing industrial chemicals, which are expected to be of low toxicity, under the REACH regulation. However, also for such uses, read-across, SARs and (Q)SAR approaches should be considered to support the interpretation of in vitro cytotoxicity data, in particular when testing substances with complex chemistry that may, although not designed to do so, exert effects on specific molecular targets leading to acute systemic toxicity. The limitations described in the OECD guidance document Nr. 129 on the use of the 3T3 NRU assay for determining the starting dose (OECD 2010) may moreover serve as a starting point for a further description of the possible limitations of the assay. However, the ESAC also acknowledges that it is at present not sufficiently clear whether and to which extent substances acting via specific modes of action may also exhibit effects on basal cellular physiology and, hence, to which extent the 3T3 NRU may therefore be applicable also to such substances. Application of the 2000 mg / kg threshold to the summary data of the previous NICEATM/EURL ECVAM validation study data which contained an appreciable number of pharmaceuticals and pesticides suggests that the applicability domain of the assay may extend to cover substances that have specific modes of action.

In summary, ESAC considers that the 3T3 NRU assay appears useful within a testing strategy for identification of potential negatives (i.e. substances with LD50 > 2000 mg/kg b.w.), in particular in the context of assessing industrial chemicals e.g. under the REACH regulation. Nevertheless the ESAC has some reservations regarding the immediate feasibility of such use due to, currently, the absence of an agreed conceptual framework for such a strategy and, in particular, due to the need to define

(a) exclusion criteria for chemicals that should not be tested in the 3T3 assay (or whose data, if tested, should be treated with caution) and

(b) possible tools for generating information / data to satisfy such exclusion criteria.

To support the further definition of the assay's applicability domain, which may be larger than currently assumed, and in order to define suitable exclusion criteria of chemicals when using the assay within a testing strategy, the ESAC suggests that additional chemicals be tested using the simplified or automated protocol variants.

Such testing should include (a) chemicals of high toxicity, which were not tested to a sufficient extent during the ESAC follow-up study and which would improve the appraisal of the likelihood of possible False Negatives (FNs) of high concern and (b) substances designed to act via known and specific modes of action (e.g. pesticides and pharmaceuticals).

1. Mandate of the ESAC

1.1 Summary of the mandate

On its 34th meeting in March 2011, EURL ECVAM mandated ESAC (EC/EURL ECVAM 2011; Annex 2) with the scientific peer review of an ECVAM-coordinated follow-up study on the 3T3 Neutral Red Uptake (NRU) assay. Building on previous work, notably the joint NICEATM/EURL ECVAM study acute oral toxicity (NIH 2006), the EURL ECVAM follow-up study had addressed the assay's capacity to reliably identify substances not requiring classification for acute oral systemic toxicity (i.e. 'negatives') on the basis of the 2 000 mg/kg b.w. threshold of the EU through the Classification, Labelling and Packaging (CLP) regulation (EU 2008) implementing the Globally Harmonised System (GHS) for classification.

Prior to the meeting, EURL ECVAM had forwarded to ESAC a formal request for scientific peer review, outlining the study's scientific background, regulatory rationale, design and key conclusions and detailing ESAC's mandate, i.e. the review's intended objective, questions, timelines and deliverables (EC/EURL ECVAM 2011; Annex 2). Briefly, the review's aim was to assess whether study design and conduct were appropriate in view of the study's objective and whether the conclusions drawn in the Validation Study Report (VSR) were substantiated by the study results and plausible in the context of other relevant information. Notably, ESAC was requested to assess the possible use of the 3T3 NRU assay within an Integrated Testing Strategy (ITS) to screen out potential negatives.

Following discussion of this request, the ESAC decided to set up a Working Group (WG) of six experts to prepare a detailed scientific peer review of the EURL ECVAM Validation Study Report (VSR) and other relevant information. Three ESAC members and three external experts were appointed following proposals by ESAC, EURL ECVAM and nominations by ECVAM's ICATM partner organisations (see Annex 1 of the ESAC WG report). The ESAC WG met once in person at EURL ECVAM in Ispra from 12-14 September 2011 and communicated through telephone conferences and written procedure.

The ESAC WG reviewed the EURL ECVAM follow-up study report (=Validation Study Report, VSR) and took also other existing information into consideration, e.g. the previous NICEATM/EURL ECVAM validation study (NIH, 2006) and an EURL ECVAM analysis of the percentage of substances not requiring classification for acute oral systemic toxicity (negatives) amongst industrial chemicals that have been notified prior to marketing (Bulgheroni et al., 2009). The ESAC discussed draft WG report and main findings at the 35th ESAC meeting in October 2011 and agreed on key observations and conclusions of its opinion. The ESAC secretariat subsequently drafted this opinion and provided it to the ESAC in advance to its 36th meeting on 20 March 2012 during which it was discussed and endorsed by the ESAC plenary.

1.2 General objective of the advice to be given by ESAC

The opinion of ESAC should support EURL ECVAM with respect to the development of further Recommendations regarding the ability of the 3T3 NRU test method to correctly identify substances not requiring classification for acute oral toxicity (LD50 > 2000 mg/kg b.w.) according to the EU CLP system implementing the Globally Harmonised System (GHS). In particular, the ESAC review should provide advice on the potential usefulness of the test method in a tiered testing approach or an Integrated Testing Strategy (ITS) for acute oral toxicity testing. To this end, the ESAC opinion should, if appropriate, point to possible gaps between the study findings and the possible standardized application of the 3T3 NRU assay for screening out negatives and suggest necessary follow-up work. The specific review questions related to this mandate are listed in section 4.2 of Annex 2, the EURL ECVAM request form for ESAC advice (EC-ECVAM 2011; Annex 2).

1.3 Background to the ESAC Mandate

1.3.1 Background to the study

The EURL ECVAM follow-up study, performed from 2008 to 2011, was planned on the background of a substantial amount of previous studies assessing the possibility to use toxic effects on cells in vitro for extrapolating on acute systemic toxicity in vivo (i.e. classification and labelling purposes). A significant correlation between in vitro cytotoxicity and in vivo toxicity has been first suggested from pharmaceutical studies. Testing potential anti-cancer drugs in vitro and in vivo, Eagle and co-workers observed a positive correlation between cytotoxicity (IC50) data and acute toxicity data (intravenous LD50 in mice) (Eagle & Fowley, 1958 a,b).

The Multicentre Evaluation of In Vitro Cytotoxicity (MEIC) programme initiated in 1983 involved 59 laboratories worldwide that used in-house protocols to test 50 reference chemicals selected by the Swedish Poison Control Centre on the basis of available reference data, including data on human toxicity and kinetics to allow the approximation of 50% lethal blood concentration (LC50) relationships that could be compared to in vitro cytotoxicity data obtained with these reference chemicals. While the MEIC programme showed that in vitro methods used in the study predicted human acute oral lethality better than did mouse and rat in vivo LD50 data, it was also recognised that other important mechanisms of toxicity exist that are typically not modelled in simple cytotoxicity tests. The MEIC programme's results have been published in a series of papers (for bibliographic references see Gennari et al., 2004).

Also in the 1980s Halle and colleagues began to collect available data on the in vitro / in vivo correlation and assembled a database of substances with acute toxicity data from in vivo and accompanying in vitro cytotoxicity data. They suggested the use of in vitro cytotoxicity as a possible predictor of acute oral toxicity (Halle & Goeres, 1988; Halle 1998). This database (called the "Registry of Cytotoxicity", RC) contains data from regulatory agencies in Germany and the US (notably the US RTECS database = Registry of Toxic Effects of Chemical Substances of the National Institute for Occupational Safety and Health, NIOSH, US) as well as from industry. All data were collected according to defined acceptance criteria. The database is currently maintained by the Federal Institute for Risk Assessment in Germany and contains about 540 chemicals with combined in vivo / in vitro data. Importantly, on the basis of the data available, a regression formula was determined empirically (millimole regression) for predicting, on the basis of IC50 values, (a) appropriate starting doses for acute oral toxicity tests in vivo and (b) LD50 values for eventual classification and labelling purposes. The regression formula is as follows: log LD50 (mmol/kg) = 0.435 x log IC50 (mM) + 0.625.

On the basis of these activities and the observation that cytotoxicity tests alone were unlikely to be capable of fully replacing the animal test, a joint international NICEATM/EURL ECVAM validation study (NIH, 2006) was conducted from 2000 to 2006 using primary normal human epidermal keratinocytes and BALB/c 3T3 mouse fibroblasts to evaluate whether cell viability/cytotoxicity (measured through NRU uptake) could be used to predict the starting dose for systemic acute oral toxicity test methods in vivo. The RC-derived regression model (either expressed in millimole or weight) was used to assess the predictive capacity of the tests on the basis of 72 substances, mostly (58/72=81%) taken from the RC database. Moreover, the study assessed the accuracy of the cytotoxicity assays to estimate rodent oral LD50 values across the six categories for acute oral toxicity of the Globally Harmonized System (GHS) (five toxicity classes plus one category for unclassified substance, considered negatives). The study concluded that the two assays were reliable and could be used in a "weight-of-evidence approach" to determine the starting dose for acute oral in vivo toxicity protocols (Up and down procedure and acute toxicity class method). The study showed that while the overall accuracy of the 3T3 NRU test method for correctly predicting each of the GHS acute oral toxicity classification categories was low (around 30%). However, substances falling in the low toxic

category (GHS class 4: 300 < LD50 • 2.000 mg/kg b.w.) were predicted with appreciable accuracy of 81%.

1.3.2 Specifics of this study

Building in particular on the previous prospective NICEATM/EURL ECVAM prospective validation study on the use of the 3T3 NRU cytotoxicity assay for acute toxicity testing, the EURL ECVAM follow-up study deviates to some extent from the typical design of a full prospective validation exercise (see schematic depiction in Figure 1). These specific are briefly recapitulated below:

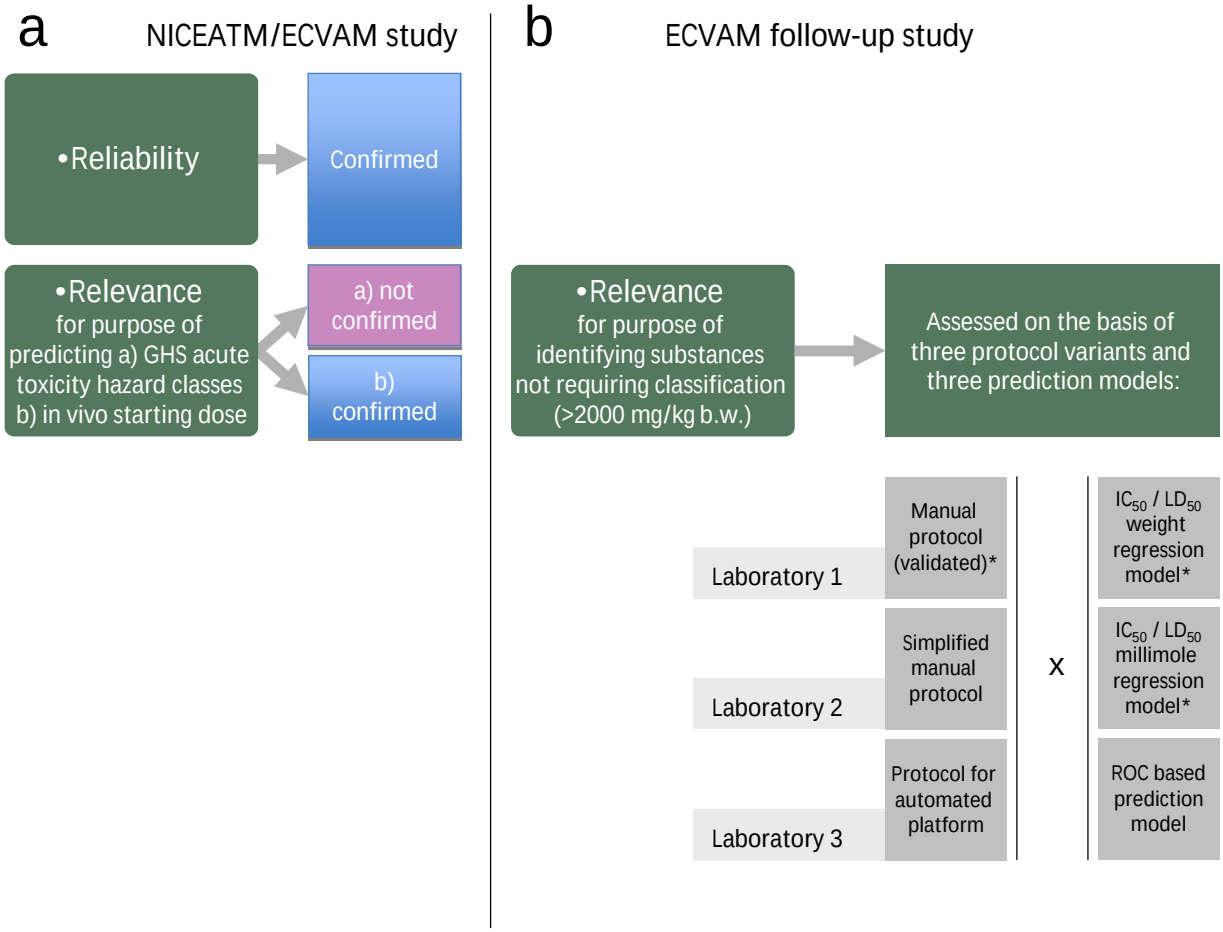
Focus on predictive capacity assessment: On the basis of the work briefly summarised in 1.3.1 and the fact that the assay protocol had been validated with regard to its reproducibility during the NICEATM/EURL ECVAM study, the EURL ECVAM follow-up study did not assess protocol reliability (within-/between-laboratory reproducibility, transferability) but instead focused on a specific aspect of the test method's predictive capacity, that is to identify specifically substances considered negatives for acute systemic oral toxicity and not requiring classification and labelling according to the legal provisions in Europe, the EU Classification, Labelling and Packaging (CLP) regulation implementing the UN Globally Harmonised System (GHS). Substances considered negative in the EU are those with LD50 values that are greater than or equal to 2 000 mg per kg body weight (b.w.). In total 56 substances specifically selected for this study were assessed. Importantly the substances were not contained in the RC, on the basis of which the prediction model (regression line) has been determined.

Use of negative test results only: The principal aim of the follow-up study was to assess whether the 3T3 NRU can be used in a way to accurately identify negatives in order avoid that all substances that require empirical testing data for acute toxicity are being tested in vivo as is currently done. Therefore, although the test provides in principle a binary classification outcome (positive or negative), only negative test results are to be considered. Thus, should the 3T3 NRU assay be found useful as a first filter within a testing strategy, in vivo testing could focus only on the confirmation of positives results from the 3T3 NRU assay and, if indeed positive in vivo, allow classification according to the four toxicity classes of EU CLP. The scientific / statistical rationale behind this proposal was the high sensitivity of the 3T3 NRU assay. Briefly, the high sensitivity means that a very high percentage of positives are indeed correctly identified. Therefore, negative test results in the 3T3 NRU are very likely to represent true negatives (positives have been sieved out). Negative test results therefore should be considered as a relevant result of the assay. In contrast, positive results cannot be considered: although containing the majority of True Positives, they also contain a high rate of False Positives (a consequence of the high sensitivity/low specificity) and in any case do not allow yet classification according to the oral acute toxicity classes. Clearly, considerations of applicability domain / limitations influence the use of the test method beyond the mere aspect of statistically measurable performance (sensitivity) on the basis of a validation testing set (see section 2. Detailed Opinion of the ESAC).

Two additional protocol variants in view of using the test for screening purposes: As the objective of the study was to assess the possible use of the 3T3 NRU assay within a tiered testing approach or an Integrated Testing Strategy (ITS) protocol variants adapted for screening approaches were tested in addition to the manual protocol previously shown reproducible: The additional protocol variants were 1) an abbreviated version of the validated manual protocol targeted at resolving toxicities around 2000 mg/kg b.w. and 2) a protocol modified for use on an automated platform. This additional testing was intended to provide information on the extent to which the original protocol is amenable to simplification (protocol variant 1), and automation (protocol variant 2). The three protocol variants were assessed by one laboratory each (manual protocol: HSL laboratory; simplified protocol: IIVS laboratory; automated protocol: JRC laboratory).

Three prediction models: In addition, three prediction models were analysed: the two regression models that had been used in the previous NICEATM/EURL ECVAM study (NIH 2006) based on the RC regression model (Halle 1998, 2003) and expressed either on the basis of concentration (millimole regression) or weight (weight regression). Third a threshold prediction model had been defined for the purpose of the study on the basis of the RC chemical data set and by using a Receiver Operating Characteristic (ROC) curve allowing the analysis of prediction thresholds optimised either for maximal sensitivity, maximal specificity or a balance of the two. Importantly, each of the three laboratories used all of three prediction models for predicting non-classified substances so that data of each of the three protocol variants were translated into concrete predictions on the absence of acute oral toxicity (i.e. LD50 > 2000 mg/kg b.w.).

Figure 1: Simplified schematic depiction of the study designs of a) the NICEATM/EURL ECVAM study assessing the ability of the 3T3 NRU assay to predict the starting dose for in vivo acute oral toxicity tests as well as the acute toxicity classes of GHS and b) of the EURL ECVAM follow-up study assessing the predictive capacity of the 3T3 NRU assay for identifying substances not requiring classification ("negatives"). a) The NICEATM/EURL ECVAM study could confirm the overall reliability of the 3T3 NRU assay, but found that, on the basis of the 72 substances tested, the assay is not capable of sufficiently accurately predicting LD50 values to allow distinguishing the GHS hazard classes. b) On the background of the established reliability of the assay, the EURL ECVAM follow-up study focused on the assay's ability to identifying selectively negatives using three protocol variants assessed in three laboratories as well three prediction models (assessed by each of the three laboratories). Protocols and prediction models labelled with an asterisk (*) are identical to those assessed in the previous NICEATM/EURL ECVAM study. In total 56 chemicals were assessed, which had been selected specifically for this study and were different from those of the NICEATM study and also different from those contained in the Registry of Cytotoxicity (RC), which had been used to establish two of the predictions models based on regression lines.



2. Detailed opinion of the ESAC

Taking into account (a) the detailed review of the ESAC WG, (b) the information made available to ESAC by EURL ECVAM including the VSR of the follow-up study and other relevant information, (c) the EURL ECVAM request for ESAC advice outlining review questions (EC/ECVAM 2011; Annex 2), the ESAC has the following opinion:

2.1 Background, regulatory and scientific rationale

The EURL ECVAM follow-up study concerned the capacity of the 3T3 NRU assay to identify specifically substances considered "negative" in terms of acute systemic oral toxicity. These are substances not requiring classification according to the legal provisions in the EU laid out in the EU Classification, Labelling and Packaging (CLP) Regulation (EU 2008) which implements the UN Globally Harmonised System (GHS). All substances with LD50 values greater than 2000 mg per kg of body weight fall into this category "non-classified".

The study was planned on the background of a substantial amount of previous work that showed that there is appreciable correlation between impaired viability of cells in vitro (measurable e.g. through vital stains or live dyes) following chemical exposure and acute (oral) toxicity in vivo (as traditionally described on the basis of population LD50 values). This correlation is believed to be due to a common causal origin i.e. the impairment of cell viability (=cytotoxicity) in vitro as well as in vivo. Importantly, the toxic effects picked up by general cell viability measurements are thought to be due to the impairment of the basal physiology and molecular biology universal to eukaryotic cells, independent of their differentiation state (e.g. as hepatic, cardiac cells, neurons, fibroblasts etc.). According to this consideration, substances can be toxic on a systemic level because they perturb basal cellular physiology/molecular biology leading to impaired functioning and viability of cells, tissues and organs, ultimately causing toxicity on a systemic level. Due to the complexity of cellular physiology, a wide variety of mechanisms are potentially targets for toxic substances, including pH regulation, integrity of plasma membrane and other membranous organelles in the cell, integrity of membrane-anchored proteins, energy metabolism etc (Gennari et al., 2004). Importantly, the measurement of basal cytotoxicity alone does not allow extrapolating on more specific modes of action expressed in an organ-specific manner (i.e. target-organ toxicity). Thus, substances exerting their toxic effects through specific mechanisms will not likely be picked up in assays based on basal physiology/cytotoxicity. In the current context, it should therefore be noted that the BALB/c 3T3 cell line is of fibroblast origin and is not expected to recapitulate organ-specific physiological mechanisms ("target-organ toxicity") that may be targeted by specific toxic substances (i.e. cardiac channels, specific hepatic proteins etc.). These considerations have consequences on the expected applicability domain and limitations of the test method.

2.2 Design and conduct

The ESAC observed that the study had been designed and conducted appropriately in view of the stated objective of using the test for identifying negatives in a tiered approach/testing strategy. However, some shortcomings with regard, in particular, to chemical selection were observed. For example, the exclusion criteria employed for selecting test items were defined rather restrictively in order to exclude risks to the study personnel. For instance CMRs (carcinogenic, mutagenic and reproductive toxicants) were excluded. This may have led to a rather small available pool of highly toxic substances amongst the test items in the validation set. In addition, the ESAC had minor concerns regarding the derivation of reverence LD50 values through simple averaging in case several values were available for one substance. The ESAC is of the opinion that a single reference value should have been obtained by applying the median. The ESAC however acknowledges that LD50 values are inherently imprecise (Bass et al, 1982; Rowan et al., 1983). As a result, it is in any case difficult to derive “true” LD50 values (see Table

4 of the VSR, and Section 2.4 of the ESAC WG Report including Figure 2.). The ESAC furthermore acknowledges that there are no other useful reference data easily available. For instance, human data from accidental exposure are even more unreliable due the many sources of uncertainty related to the actual exposure levels. Other minor issues included that clearly outlying LD50 values had not been deleted, although this concerned only two chemicals amongst the 56 tested.

Strengths of the study include the practice of censoring and excluding values in a transparent and plausible manner. This had minor implications however on the size and completeness of the data matrices generated and used for final analysis: Although the number of chemicals is appropriate for the purpose of the study according to a short statistical analysis of the ESAC WG, censoring and excluding uncertain values reduced the size of the finally usable datasets. The ESAC is of the opinion that this reduction of available data points could have been anticipated and that the number of test substances should, ideally, have been slightly larger to accommodate such foreseeable reductions.

2.3 Study results and conclusions

Prediction model and predictive performance: In agreement with the EURL ECVAM validation study report (VSR), the ESAC is of the opinion that the original millimole regression seems to be the most useful prediction model. Independent of the protocol used, the Sensitivity (True Positive Rate) on the basis of the millimole regression was very high at about 95%, supporting the use of the assay and this particular prediction model for the identification of negatives. In contrast, predictions obtained with the weight regression model were less accurate. Due to the lack of some information on how the ROC threshold values had been derived, the ESAC could not review this third prediction model.

Applicability domain and restrictions with regard to specific mechanism of actions on biological targets: In this context, the ESAC emphasises however that this high sensitivity is obtainable only for the applicability domain established on the basis of considerations of mechanism of action that are transparently laid out in the VSR: substances with a purposefully designed mechanism of action on molecular targets (i.e. receptors, channels, enzymes) should not be tested in the 3T3 NRU assay as specific modes of action are not modelled in this fibroblast cell line. This includes for instance pharmaceuticals and pesticides. Moreover, the selection criteria applied when creating the validation testing set excluded some classes of substances without specific mechanisms of actions but containing potentially highly toxic substances (e.g. mutagenic substances), leaving the possibility that, upon inclusion of substances of higher toxicity, the sensitivity obtained might have been lower.

The ESAC acknowledges on the other hand, that there is indication that the applicability domain may extend beyond substances acting only via cytotoxicy and also apply to substances with a specific mode of action. This notion is based on the application of the 2000 mg/kg b.w. threshold to the validation summary data from the previous NICEATM/EURL ECVAM study. The 72 chemicals tested during this validation study included an appreciable percentage of pharmaceuticals and pesticides (42% and 22%, respectively; see categorisation in NIH 2006). When applying the LD50 cut-off value of 2000 mg/kg b.w. to these data, nearly all true negatives were correctly identified. Importantly, there was only one False Negative, the plant alkaloid “digoxin” see recalculation of the NICEATM/EURL ECVAM data set, Annex 2 of ESAC WG report). Thus, although the applicability domain, as laid out in the VSR, currently excludes pharmaceuticals and pesticides, it seems not sufficiently clear whether and to which extent toxic substances exerting effects via specific mechanisms of action may also cause appreciable effects via basal cytotoxicity. This may allow use of the 3T3 NRU assay on a much broader population of substances. Finally, The ESAC acknowledges in this context that the issue of

extrapolating from a limited testing set during validation on a probable applicability domain is a general difficulty of any validation study and not specific to this study.
Moreover, while the ESAC considers that the 3T3 NRU assay appears useful within a testing strategy and may have an applicability larger than currently assumed, the ESAC has some reservations regarding the immediate feasibility of such use due to, currently, the absence of an agreed conceptual framework for such a strategy and, in particular, due to the need to define
(a) exclusion criteria for chemicals that should not be tested in the 3T3 assay (or whose data, if tested, should be treated with caution) and
(b) possible tools for generating information / data to satisfy such exclusion criteria.
While some substances could be easily excluded from testing, e.g. those designed for oral uptake and acting via a specific mechanism of action (pharmaceuticals) OR acting via a specific mechanism and likely to be (at least in background doses) ingested (pesticides), such exclusion would be considerably more difficult for complex organics containing functional groups that may activate specific molecular targets although not having been designed to do so.
Protocols amenable for screening approaches: As the two non-validated protocols used in the study (one manual simplified version and one version run on an automated platform at JRC) gave comparable results to those obtained with the previously validated protocol (NICEATM-EURL ECVAM study), the ESAC considers that the test method is apparently robust and amenable to simplification / automation and may therefore be a candidate assay for high throughput screening approaches.
Potential reduction of animal testing: The ESAC agrees further with the observation presented in the VSR according to which the 3T3 NRU assay may allow reducing the number of necessary animal tests by a significant proportion. This reasoning is based on the test method's high sensitivity and the assumption that a significant number of chemicals requiring acute oral toxicity data are negatives (EURL ECVAM analysis of the proportion of notified chemicals in the EU New Chemicals Database (Bulgheroni, 2008). However, the ESAC considers that a note of caution is required with respect of the potential animal saving brought about by the 3T3 NRU test as this is dependent on the applicability domain restrictions of the test and will further depend (a) on the actual percentage of true negatives, currently not known for chemicals in the developmental phase in industry during which the test may be used (b) the provisions of an eventual Integrated Testing Strategy which at present is unavailable. However, it should be noted that in the area of acute oral toxicity testing there are currently no non-animal methods employed for generating data on the presence or absence of potential hazards. Thus, nearly irrespective of the percentage ("prevalence") of negatives amongst chemicals, any test method allowing identification of one category of substances (here negatives) will likely contribute to a reduction of animal testing.

2.4 Recommendations

The ESAC recommends the generation of further testing data to address issues of applicability domain and also possible False Negatives of higher concern (i.e. positives that are highly toxic).
The ESAC recommends that additional data be generated on the basis of the protocol variants for screening purposes (simplified or automated variant). It would appear sufficient to produce these data in a single laboratory as the present study has shown that these protocol variants showed results comparable to the previously validated manual protocol and that therefore, reproducibility does not seem to be a problem. Such additional testing should include (a) a significant number of highly toxic substances, (b) substances known to act on specific molecular targets, i.e. pharmaceuticals, pesticides, biocides etc. Such testing data will subsequently help defining, in a more precise manner, the applicability and limitations of the 3T3 NRU assay for use within a testing strategy. It should be noted that the 3T3 NRU assay (protocol and millimole prediction model) is already in use for regulatory purposes, although only to define starting

doses for the in vivo tests. Guidance document 129 of the OECD outlines the use of the 3T3 NRU assay for this purpose, also stipulating limitations concerning specific mechanism of action and metabolism. These could probably serve as a starting point for defining limitations for the 3T3 NRU assay for the purpose of generating data on the absence of toxicity within a testing strategy.

As the 3T3 NRU test allows measurement of basal cytotoxicity but is unlikely to model specific mechanisms of action, it seems best adapted to support hazard and risk assessment of substances not designed to interact with specific molecular targets and assumed a priori to be of low toxicity. Thus, the 3T3 NRU may be particularly useful for assessing, in the context of the REACH legislation, industrial chemicals not covered by the above mentioned exclusion criteria. In addition the test method may, if appropriately employed in a testing strategy / tiered approach, also be useful for non-regulatory screening purposes during substance development.

Finally, the ESAC recommends a minor amendment of the SOP of the 3T3 NRU assay concerning solubility: the approach used by the IIVS laboratory in the present study seems useful to increase potentially the solubility of substances and should be incorporated in the general SOP.

3. References

Bass R, Gunzel P, Henschler D (1982) LD502 versus acute toxicity. Arch. Toxicol. 51,183-186.

Bulgheroni A, Kinsner-Ovaskainen A, Hoffmann S, Hartung T, Prieto P (2009) Estimation of acute oral toxicity using the No Observed Adverse Effect Level (NOAEL) from the 28 day repeated dose toxicity studies in rats. Regul Toxicol Pharmacol. 53(1):16-9.

Eagle H, Foley GE (1958)a. The cytotoxicity of anti-tumor agents for normal human and animal cells in first tissue culture passage. Cancer Res. 18(9):1011-6.

Eagle H, Foley GE (1958)b. Cytotoxicity in human cell cultures as a primary screen for the detection of anti-tumor agents. Cancer Res. 18(9):1017-25.

EURL EURL ECVAM (2011): EURL ECVAM Request for ESAC advice on the ECVAM-coordinated follow-up study to assess the predictive capacity of the already validated Neutral Red Uptake Cytotoxity assay for acute oral toxicity testing.

EURL ECVAM (2010): Validation Study Report of the ECVAM-coordinated follow-up study on the predictive capacity of the 3T3 NRU assay for acute oral toxicity testing.

EURL ECVAM (2012): ESAC Working Group Peer Review Consensus Report on the ECVAM-coordinated follow-up study to assess the predictive capacity of the already validated Netural Red Uptake Cytotoxity assay for acute oral toxicity testing.

European Union (2008) REGULATION (EC) No 1272/2008 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 16 December 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing Directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006. Online:

http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:353:0001:1355:EN:PDF

Halle W, Goeres E (1988) Register der Zytotoxizit ät (IC50) in der Zellkultur und Möglichkeiten zur Abschätzung der akuten Toxizität (LD50). In: Beiträge zur Wirkstofforschung, Institut für Wirkstofforschung (ed. P. Oehme, H. Loewe & E. Goeres), 108 pp. Berlin, Germany.

Halle W (1998) Toxizitätsprüfungen in Zellkulturen für eine Vorhersage der akuten Toxizität (LD50) zur Einsparung von Tierversuchen. Schriften des Forschungszentrums Jülich: Reihe Lebenswissenschaften/Life Sciences, Vol.1 93pp. Jülich, Germany

Halle W (2003) The Registry of Cytotoxicity: Toxicity Testing in Cell Cultures to Predict Acute Toxicity (LD50) and to Reduce Testing in Animals. ATLA 31, 89-198

NIH (2006) Background Review Document (BRD): Validation of Neutral Red Uptake Test Methods NIH/In Vitro Cytotoxicity Test Methods for Estimating Acute Systemic Toxicity. Publication NO. 07-4518, November 2006.http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm

OECD (2010) Series on Testing and Assessment No. 129. Guidance Document on using cytotoxicity tests to estimate starting doses for acute oral systemic toxicity tests.

Rowan A (1983) Shortcomings of LD50-values and acute toxicity testing in animals. Acta Pharmacologica et Toxicologica, 52: 52–64.

EURL ECVAM Scientific Advisory Committee

- Dr. Nathalie ALÉPÉE
- Dr. David BASKETTER
- Dr. Neil CARMICHAEL
- Prof. Jacques R. CHRETIEN
- Prof. Lucio G. COSTA
- Dr. Rodger CURREN
- Prof. A. Wallace HAYES
- Prof. Coenraad HENDRIKSEN
- Dr. Dagmar JÍROVÁ
- Prof. Walter PFALLER
- Dr. Erwin ROGGEN
- Prof. Vera ROGIERS
- Dr. Andrea SEILER
- Prof. Kristin SCHIRMER
- Prof. Ruud A. WOUTERSEN

ESAC Working Group 3T3 NRU

- Dr. Neil CARMICHAEL (ESAC member, Chair of ESAC WG)
- Prof. Lucio COSTA (ESAC member)
- Dr. Ian DEWHURST (external expert; nominated by ESAC)
- Dr. Eugene ELMORE (external expert, nominated by NICEATM)
- Prof. Annette KOPP-SCHNEIDER (external expert, nominated by ECVAM)
- Prof. Kristin SCHIRMER (ESAC member)

Coordination

- Dr. Claudius GRIESINGER (EURL ECVAM Coordinator for ESAC peer reviews and EURL ECVAM Recommendations)
- Dr. Anita TUOMAINEN (specific support)

Annex 2 EURL ECVAM request for ESAC advice

EURL ECVAM request to ESAC for scientific advice on the ECVAM-coordinated follow-up study to assess the predictive capacity of the Neutral Red Uptake cytotoxicity assay for acute oral toxicity testing

ESAC Request 2011-02

Title page information	
Abbreviated title of ESAC request	Follow-up study to assess the PC of the Neutral Red Uptake cytotoxicity assay for acute oral toxicity testing
ESAC REQUEST Nr.	2011-02
Template used for preparing request	EP 2.01
Date of finalising request	2011-03-07
Date of submitting request to ESAC	2011-03-09
Request discussed through	Plenary discussion at ESAC 34, 22-23 March 2011
Opinion expected at (date)	ESAC 35, 4-5 October 2011
File name of this request	ESAC REQUEST 2011-02 3T3 NRU follow-up-final.doc

1. TYPE OF REQUEST

Request Type	Identify request ("YES")
ESAC Peer Review of a Prevalidation Study or Validation Study	YES
If R1)applies please specify further:	
• Prevalidation Study	NO
• Prospective Validation Study	YES This study, finished in 2010, was planned and conducted as a follow-up to the previous full prospective validation study of the 3T3 NRU cytotoxicity assay conducted by NICEATM in collaboration with EURL ECVAM and finalised in 2005. The study was designed to complement the information on predictive capacity of the 3T3 NRU assay for the specific purpose of identifying substances that do not need to be labelled for acute oral toxicity according to the EU CLP regulation (i.e. substances with LD_{50} doses above the limit dose of 2000 mg/kg body weight). Importantly, being a follow-up and complement exercise according to ECVAM's modular approach, the study deviates to some extent from the typical design of a full prospective validation study (e.g. no transferability/reproducibility assessment as these have been addressed in the previous study).
• Retrospective Validation Study	NO
• Validation Study based on Performance Standards	NO
Scientific Advice on a test method submitted to EURL ECVAM for validation (e.g. the test method's biological relevance etc.)	NO
Other Scientific Advice (e.g. on test methods, their use; on technical issues such as cell culturing, stem cells, definition of performance standards etc.)	NO

2. TITLE OF STUDY OR PROJECT FOR WHICH SCIENTIFIC ADVICE OF THE ESAC IS REQUESTED

Follow-up study on the predictive capacity of the 3T3 Neutral Red Uptake cytotoxicity assay to correctly identify substances not classified for acute oral toxicity under the EU CLP system (LD_{50} > 2000 mg/kg).

3. BRIEF DESCRIPTION OF THE STUDY OR PROJECT

3.1 Summary of the follow-up study

This follow-up study was conducted as a complement to the previous full prospective validation study of the 3T3 Neutral Red Uptake (NRU) cytotoxicity assay (=the "3T3 NRU assay") conducted by NICEATM/ICCVAM in collaboration with ECVAM. The test exploits the correlation between the systemic toxicity (i.e. acute oral toxicity) of substances and their cytotoxicity exerted on 3T3 cells. Cytotoxicity is measured as reduction of uptake of the vital dye 'neutral red', which accumulates in lysosomes of healthy cells.

As a follow-up, the study deviates to some extent from the manner in which a full prospective validation exercise is typically conducted. The study was designed to specifically assess whether the 3T3 NRU assay is able to discriminate classified chemicals from non-classified ones (i.e. those beyond the limit dose of 2000mg/kg according to the EU CLP regulation implementing UN GHS). Thus, the study was intended to provide additional information on predictive capacity of the 3T3 NRU assay for this specific purpose, without addressing reproducibility/transferability of the protocol which had been previously demonstrated in the NICEATM/EURL ECVAM validation study.

To assess the capacity of the 3T3 NRU to correctly identify chemicals not requiring classification, 56 test items (a sufficient number to analyse dichotomous classifications) with good in vivo reference data were selected and tested in one laboratory using the already validated protocol. In addition, the same chemicals were tested in two more laboratories using slight modifications of the protocol. These variations were 1) an abbreviated version of the validated protocol, and 2) a protocol modified for use on an automated platform. This additional testing was intended to provide information on the extent to which the original protocol is amenable to simplification (protocol variant 1), and automation (protocol variant 2).

The testing data of the validated protocol show that the 3T3 NRU identifies true positives with a sensitivity of 94%. Since the 3T3 NRU is able to correctly identify most positives, negative test results in the 3T3 NRU are very likely to represent either true negatives (non-classified chemicals) or false positives. In contrast the rate of false negatives is low. This is reflected by the high NPV (negative predictive value) of 92%. Therefore, the 3T3 NRU may be appropriate for identifying negatives as a first screening step in a tiered testing approach involving subsequent in vivo testing to 1) further categorise chemicals with positive results, 2) to identify false positive results of the 3T3 NRU (low specificity of 42%) and 3) to test, in specific cases where there is additional weight of evidence information, negative substances for confirmation. The two protocol variants gave similar predictive values suggesting that also these variants of the validated 3T3 NRU protocol may be used for the screening of non-classified chemicals according to EU CLP within a tiered testing strategy.

3.2 Detailed background

Several international projects have studied the possibility of using in vitro methods to predict acute oral toxicity.

The first of these studies was the Multicentre Evaluation of In Vitro Cytotoxicity (MEIC) programme. It showed that in vitro methods used in the study predicted human acute oral lethality better than did mouse and rat in vivo LD_{50} data.

In a second study based on information of the Registry of Cytotoxicity (RC, a database for rodent acute oral LD_{50} values and in vitro IC_{50} values), over 70% of the substances tested in vitro were able to predict the rodent acute oral lethality.

Third, the international NICEATM/EURL ECVAM validation study (the In Vitro Basal Cytotoxicity Validation Study finished in 2005) used a human-derived cell model (primary normal human

epidermal keratinocytes) and a mouse cell model (BALB/c 3T3 mouse fibroblasts) to evaluate the usefulness and limitations of the in vitro basal cytotoxicity test methods based on measuring cell viability trough neutral red uptake (NRU) for predicting starting doses for systemic (i.e. in vivo) acute oral toxicity test methods. In addition, this validation study assessed the accuracy of the two basal cytotoxicity test methods to estimate rodent oral LD_{50} values across the five categories of the Globally Harmonized System (GHS) for acute oral toxicity as well as unclassified toxicities. The study concluded that the two NRU test methods could be used in a weight-of-evidence approach to determine the starting dose for acute oral in vivo toxicity protocols. The validation study also showed that the overall accuracy of the 3T3 NRU test method for correctly predicting each of the GSH acute oral toxicity classification categories was low (around 30%), however, substances falling in the GHS 4 category (i.e. 300 < LD_{50} • 2.000 mg/kg) were predicted better, with 81% accuracy.
Taken together, the results of MEIC, the RC, and the NICEATM/ECVAM international validation study have all shown a correlation of around 60-70% between in vitro cytotoxicity data and oral rodent LD_{50} values. These studies indicated that the in vitro methods are able to predict low systemic toxicity with much greater precision than high systemic toxicity, suggesting the potential usefulness of these methods for identifying chemicals not requiring classification.

3.3 Purpose of the study
This follow-up study was initiated in 2008 by ECVAM and was finalised in October 2010. The aim of this study was to further explore, on the basis of the previous validation study, whether the predictive capacity (e.g. sensitivity, specificity, concordance) of the 3T3/NRU cytotoxicity assay is sufficient to correctly distinguish chemicals not requiring classification for acute oral toxicity according to provisions of the EU CLP regulation (i.e. LD_{50} > 2000 mg/kg b.w.) from those that require classification (LD_{50} • 2000 mg/kg b.w.). The scientific and regulatory rationale embedded in study's objective was to assess whether the 3T3 NRU assay could be used as the first step of a tiered approach to identify unclassified chemicals so that subsequent testing in vivo would focus on confirmatory testing to classify positives according to the 4 classified classes of EU CLP and identify substances with positive test results in the 3T3 NRU that are actual negative (=3T3 NRU false positives).
The study used the test method protocol validated in the NICEATM/ECVAM validation study. In addition, two protocol modifications were assessed: one version of the 3T3/NRU protocol adapted to an automated platform and an abbreviated version of the validated protocol that was targeted at resolving acute oral toxicities around the 2000 mg/kg cut-off value. The aim of this additional testing was to assess whether a simplified version and a version adapted for automated testing would generate similar data on the basis of the 56 test chemicals selected and to assess, therefore, to which extent these variants of the validated protocol may be used for purposes of identifying negatives (LD_{50} > 2000 mg/kg b.w.).

3.4 Organisation of the study
The study was coordinated and managed by a Validation Management Team composed of two ECVAM staff members. Although testing was performed in three laboratories, the core validation exercise (aiming at more detailed information on predictive capacity) concerned only laboratory Nr. 1 which worked with the validated protocol. Laboratories 2 and 3 produced additional data on the basis of two protocol variants supporting a comparative analysis of protocol performance. The laboratories were:
Health and Safety Laboratory (HSL), UK (under ECVAM sponsored contract) using the already validated manual test method protocol
JRC (IHCP), Italy using the automated version of the test method protocol
IIVS, US (sponsored by IIVS and PETA, People for the Ethical Treatment of Animals) using the abbreviated test method protocol

A set of 56 coded industrial chemicals (including cosmetic ingredients) were tested using each test method protocol. The chemicals were purchased from Sigma-Aldrich (Italy) and coded by ECVAM. The distribution of chemicals and respective material safety data sheets were done by Sigma-Aldrich Germany (for the two European laboratories) and Sigma-Italy for the laboratory in the US. The data from blind testing were de-coded and analysed independently by ECVAM.

3.5 Results and conclusions

The results of all three protocol variants show that the 3T3/NRU assay has high sensitivity (92-96%) and high negative predictive value (86-92%). This indicates that compounds identified as negatives by the method (40% - 44%) will most likely be correctly categorised as unclassified (LD50 > 2000 mg/kg b.w.). Therefore, if the above proposed tiered strategy is applied, negatives may not be required to be tested in subsequent confirmatory in vivo testing. Positives of the 3T3 NRU however would require confirmatory in vivo testing on the basis of a starting dose approach as validated in the NICEATM-ECVAM validation study.

A recent analysis of the New Chemicals Database showed that over 85% of new industrial chemicals do not require classification for acute oral toxicity according to EU CLP (LD50 > 2000 mg/kg b.w.). With the 3T3/NRU method, which was demonstrated of being able to correctly identify about 42% of all true negatives, a testing strategy could be developed, limiting animal testing to only those substances identified as "classified" by the 3T3/NRU assay.

References

Clemedson et al. 1996. MEIC evaluation of acute systemic toxicity. Alt. Lab. Anim.24: 249-272

Ekwall B, Clemedson C, Craaford B, Ekwall B, Hallander S, Walum E, Bondesson I. 1998a. MEIC evaluation of acute systemic toxicity. Part V. Rodent and human toxicity data for the 50 reference chemicals. Alt. Lab. Anim. 26 (suppl 2):571-616.

Ekwall B, Clemedson C, Craaford B, Ekwall B, Hallander S, Walum E, Bondesson I. 1998b. MEIC evaluation of acute systemic toxicity. Part VI. The prediction of human toxicity by rodent LD50 values and results from 61 in vitro methods. Altern Lab Anim 26 (suppl 2):617- 658.

Ekwall B, Clemedson C, Ekwall B, Ring P, Romert L. 1999. EDIT: A new international multicentre programme to develop and evaluate batteries on in vitro tests for acute and chronic systemic toxicity. Alt. Lab. Anim. 27:339-349

Ekwall B, Ekwall B, Sjostrom M. 2000. MEIC evaluation of acute systemic toxicity: Part VIII. Multivariate partial least squares evaluation, including the selection of a battery cell line tests with a good prediction of human acute lethal peak blood concentrations for 50 chemicals. Alt. Lab. Anim. 28 (suppl 1):201-34.

Halle W. 2003. The Registry of Cytotoxicity: Toxicity Tesing in Cell Cultures to predict Acute Toxicity (LD50) and to reduce Tesing in Animals. Alt. Lab. Anim. 31: 89-198.

Anon, 2006. Background Review Document (BRD): Validation of Neutral Red Uptake Test Methods NIH /In Vitro Cytotoxicity Test Methods for Estimating Acute Oral Systemic Toxicity. Publication No. 07-4518, November 2006, http://iccvam.niehs.nih.gov/methods/acutetox/inv_nru_brd.htm.

EU 2008. Regulation (EC) No 1272/2008 on classification, labelling and packaging of substances and mixtures, amending and repealing directives 67/548/EEC and 1999/45/EC, and amending Regulation (EC) No 1907/2006 (REACH).

Bulgheroni, A., Kinsner-Ovaskainen, A., Hoffmann, S., Hartung, T., Prieto, P., 2009. Estimation of acute oral toxicity using the No Adverse Effect Level (NOAEL) from the 28-day repeated dose toxicity studies in rats. Regulatory Toxicology and Regulatory Toxicology and Pharmacology 53, 16–19.

Hartung, T., Bremer, S., Casati, S., Coecke, S., Corvi, R., Fortaner, S., Gribaldo, L., Halder, M., Hoffmann, S., Roi, A.J., Prieto, P., Sabbioni, E., Scott, L., Worth, A., Zuang, V., 2004. A modular approach to the ECVAM principles on test validity. Alt. Lab. Anim. 32:467–472.

4. OBJECTIVES, QUESTIONS, TIMELINES

4.1 Objective

Objective Why does ECVAM require advice on the current issue?	The opinion of ESAC should support ECVAM with respect to the development of further Recommendations regarding the ability of the 3T3/NRU test method to correctly identify substances not requiring classification for acute oral toxicity under the EU CLP system (LD50 > 2000 mg/kg b.w.) and the use of the test method in a tiered testing approach for acute oral toxicity testing.

4.2 Question(s) to be addressed

Questions What are the questions and issues that should be addressed in view of achieving the objective of the advice?	1) DESIGN & CONDUCT OF STUDY: The ESAC is requested to review whether the validation study was conducted appropriately in view of the objective of the study, i.e. to assess the ability of the 3T3/NRU test method to correctly identify substances not requiring classification for acute oral toxicity under the EU CLP system (LD50 > 2000 mg/kg b.w.). In particular the following issues should be addressed: (a) Clarity of the definition of the study objective. (b) Appropriateness of the study design in view of study objective, inter alia: o Were the criteria for chemical selection appropriate? o Is the toxicity range of the selected chemicals appropriate for the purpose of the study (i.e. analysis of the ability to distinguish at the 2000mg/Kg b.w. threshold)? o In case of gaps (chemical class etc.) – are these justified? o Is the number of chemicals sufficient? o Is the number of laboratories sufficient? (c) Appropriateness of the study execution (e.g. were there pre-defined acceptance criteria, were these respected? How were exceptions / deviations handled, e.g. censoring of values, retesting etc?). (d) Appropriateness of the statistical analysis as used in the protocols and for analysing predictive capacity. 2) CONCLUSIONS OF STUDY: The ESAC is requested to assess whether the conclusions, as presented in the Validation Study Report (VSR), are substantiated by the information generated during validation and are plausible with respect to existing information and current views (e.g. literature). In particular: (a) Do the data on the basis of these chemicals provide new information on applicability and possible limitations (in addition to the original information available upon completion of the original ICCAM/ECVAM study)? (b) Are the conclusions on predictive capacity justified and plausible with respect to existing information (c) Is the information on the two protocol variants (abbreviated and automated version) sufficient in view of supporting their standardized use alongside the already validated protocol? (d) Are there possible gaps between study design and study conclusions which remain to be addressed in view of the suggested conclusions / use (see also point 3)? 3) SUGGESTED USE OF THE TEST METHOD: The ESAC is requested to review the suggested use of the validated method within a strategy to identify only unclassified chemicals (LD50 > 2000 mg/kg b.w.) as proposed by the Validation Management Team. (a) Is the suggested use of the test method, based on the information generated in the Validation Study, plausible and scientifically justified? (b) Is there additional information required (i.e. are there gaps) to be able to conclude on the plausibility of the suggested use?

4.3 Timelines

Timelines concerning this request When does ECVAM require the advice?	Timeline	Indication
	Finalised ESAC Opinion required by:	ESAC 35, 4-5 October 2011
	Request to be presented to ESAC by written procedure (e.g. due to urgency) prior to the next ESAC	NO
	Request to be presented to ESAC at ESAC plenary meeting	ESAC 34, 22-23 March 2011

5. ECVAM PROPOSALS ON HOW TO ADDRESS THE REQUEST WITHIN ESAC

5.1 ECVAM proposal regarding request-related structures required

Specific structures required within ESAC to address the request Does the advice require an ESAC working group, an ESAC rapporteur etc.?	Structure(s) required	Required according to ECVAM? (YES/NO)
	ESAC Rapporteur	NO
	ESAC Working Group	YES
	Invited Experts	NO
	Ad S3: If yes – list names and affiliations of suggested experts to be invited and specify whether these are member of the EEP	
	If other than above (S1-S3):	

5.2 Deliverables as proposed by ECVAM

Deliverables What deliverables (other than the ESAC opinion) are required for addressing the request?	Title of deliverable other than ESAC opinion	Required? (YES/NO)
	ESAC Rapporteur Report and draft opinion	NO
	ESAC Peer Review Report and draft opinion	YES
	If other than above (D1-D2):	

6. LIST OF DOCUMENTS TO BE MADE AVAILABLE TO THE ESAC

Count	Description of document	Available (YES/NO)	File name
1	Final Study Report	yes	Final 3T3 NRU study report_March 2011
2	Study protocol HSL	yes	Annex 1_Study protocol of HSL
3	Study protocol JRC	yes	Annex 2_Study protocol of JRC
4	Study protocol IIVS	yes	Annex 3_Study protocol of IIVS
5	Solubility protocol	yes	Annex 4_ Solubility protocol
6	Seidle et al._2010_Cross sector review drivers and available 3Rs approaches acute toxicity testing	yes	2010_Seidel et al._ Toxicological Sciences
7	Creton et al._2010_Acute toxicity testing of chemicals—Opportunities to avoid redundant testing and use alternative approaches	yes	2010_Creton et al., Critical Reviews in Toxicology
8	Bulgheroni et al._2009_Estimation of acute oral toxicity using NOAEL	yes	2009_Bulgheroni et al._Toxicology In Vitro

7. TERMS OF REFERENCE OF THE ESAC WORKING GROUP

7.1 Establishment of the ESAC Working Group

During its 34th meeting on 22./23. March the ESAC plenary unanimously decided to establish an ESAC Working Group charged with the detailed scientific review of the ECVAM follow-up study on the predictive capacity of the validated 3T3 NRU assay for acute toxicity testing.

7.2 Title of the ESAC Working Group

Full title:
ESAC Working Group for the detailed scientific peer review of the ECVAM follow-up study of the 3T3 NRU assay for acute toxicity testing
Abbreviated title:
ESAC WG 3T3 NRU

7.3 Mandate of the ESAC WG

The EWG is requested to conduct a scientific review of the ECVAM-conducted follow-up study concerning the predictive capacity of the 3T3 NRU assay. The review needs to address the questions put forward to ESAC by ECVAM and the more detailed questions developed by the ESAC members of the ESAC WG in collaboration with the ESAC Chair, Vice Chair and Secretariat.
The review should focus on the appropriateness of design and conduct of the study in view of the study objective and should provide an appraisal to which extent the conclusions drawn in the Validation Study Report are substantiated by the information generated during the study and how the information generated relates to the scientific background available.

7.4 Deliverable of the ESAC WG

The ESAC WG is requested to deliver to the chair of the ESAC and the ESAC Secretariat a detailed ESAC Working Group Report outlining its analyses and conclusions. A reporting template has been appended (Appendix 1) intended to facilitate the drafting of the report.
The conclusions drawn in the report should be based preferably on consensus. If no consensus can be achieved, the report should clearly outline the differences in the appraisals and provide appropriate scientific justifications.

7.5 Proposed timelines of the ESAC WG

The Secretariat has proposed timelines which should be agreed upon during the first Teleconference (Item 1 in the table):

Item	Proposed date/time	Action	Deliverable
1	Mid April	Teleconference to discuss/decide 1. the list of proposed external (non-ESAC) experts for the ESAC WG 3T3 NRU 2. the more detailed questions to put forward to the ESAC WG	1. List with 3 preferred options (3 external experts + 3 ESAC members = 6 experts in total) 2. Consolidated list of questions
2	Mid April	Both deliverables of item 1 to go to the ESAC for approval / amendment	Amended deliverables as listed under item 1 (if appropriate)
3	Kick-off teleconference, second week of May (9. May – 13. May)	Discuss the organisation of review, distribution of work (if feasible). Discuss the studies. Agree on the meeting date and further timelines.	Minutes and agreed meeting date/timelines, work organisation.

7.6 Questions which should be addressed by the ESAC WG

The ESAC WG is requested to address the three questions posed to the ESAC which have been broken down further in more specific questions by the ESAC chair, the chair of the ESAC WG and the Secretariat (see section 4.2).

When preparing the final ESAC WG report to address these questions, the ESAC WG is requested to use a pre-defined reporting template. This template (see appendix 1) follows ECVAM's modular approach and addresses to which extent the standard information requirements have been addressed by the study. The template allows moreover for addressing the issues specific studies outlined in section 4.2. The Secretariat will provide guidance if necessary.

APPENDIX 1 REPORTING STRUCTURE FOR THE ESAC WG REPORT

The following suggested structure follows the ECVAM information requirements ("modules") for scientific review following validation and allows at the same time for the description of the analysis and conclusions concerning more specific questions. A template has been created on the basis of the structure below and this template will be made available to the ESAC.

The template can be used for various types of validation studies (e.g. prospective full studies, retrospective studies, performance-based studies and prevalidation studies). Depending on the study type and the objective of the study, not all sections may be applicable. However, for reasons of consistency and to clearly identify which information requirements have not been sufficiently addressed by a specific study, this template is uniformly used for the evaluation of validation studies.

1. Data collection
    - 1.1 Information / data sources used
    - 1.2 Search strategy
    - 1.3 Selection criteria applied to the available information

2. Study objective and design
    - 2.1 Clarity of the definition of the study objective
    - 2.2 Analysis of the scientific rationale provided
    - 2.3 Analysis of the regulatory rationale provided
    - 2.4 Appropriateness of the study design
    - 2.5 Appropriateness of the statistical evaluation

3. Test definition (Module 1)
    - 3.1 Quality and completeness of the overall test definition
    - 3.2 Quality of the background provided concerning the purpose of the test method
    - 3.3 Quality of the documentation and completeness of (a) standardised protocols (SOPs) and (b) prediction models

4. Data quality
    - 4.1 Overall quality of the evaluated data
    - 4.2 Sufficiency of the evaluated data in view of the study objective
    - 4.3 Quality of the reference data for evaluating reliability and relevance[2]
5. Test materials
    - 5.1 Sufficiency of the number of evaluated test items in view of the study objective
    - 5.2 Representativeness of the test items with respect to applicability

6. Within-laboratory reproducibility (Module 2)
    - 6.1 Assessment of repeatability and reproducibility in the same laboratory
    - 6.2 Conclusion on within-laboratory reproducibility as assessed by the study

7. Transferability (Module 3)
    - 7.1 Quality of design and analysis of the transfer phase

[2] OECD guidance document Nr. 34 on validation defines relevance as follows: "Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It sis the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of accuracy (concordance) of a test method."

7.2 Conclusion on transferability to a second laboratory/other laboratories as assessed by the study

8. Between-laboratory reproducibility (Module 4)
   - 8.1 Assessment of reproducibility in different laboratories
   - 8.2 Conclusion on reproducibility as assessed by the study

9. Predictive capacity (Module 5)
   - 9.1 Adequacy of the assessment of the predictive capacity in view of the purpose
   - 9.2 Overall relevance (biological relevance and accuracy) of the test method in view of the purpose

10. Applicability domain (Module 6)
   - 10.1 Appropriateness of study design to conclude on applicability domain, limitations and exclusions
   - 10.2 Quality of the description of applicability domain, limitations, exclusions

11. Performance standards (Module 7)
   - 11.1 Adequacy of the proposed Essential Test Method Components
   - 11.2 Adequacy of the Reference Chemicals
   - 11.3. Adequacy of the defined Accuracy Values

12. Readiness for standardised use
   - 12.1 Assessment of the readiness for regulatory purposes
   - 12.2. Assessment of the readiness for other uses (in house screening etc.)
   - 12.3 Critical aspects impacting on standardized use
   - 12.4 Gap analysis

13. Other considerations

14. Conclusions on the study
   - 14.1 Summary of the results and conclusions of the study
   - 14.2 Extent to which conclusions are justified by the study results alone
   - 14.3 Extent to which conclusions are plausible in the context of existing information

15. Recommendations
   - 15.1 General Recommendations concerning the study
   - 15.2 Recommendations concerning the test method (test system, protocol, prediction model)

16. References

17. Annexes

END OF EURL ECVAM RECOMMENDATION

European Commission
EUR 25946 – Joint Research Centre – Institute for Health and Consumer Protection

Title: EUR 25946 - EURL ECVAM Recommendation on the 3T3 Neutral Red Uptake Cytotoxicity Assay for Acute Oral Toxicity Testing

Luxembourg: Publications Office of the European Union

2013 – 47 pp. – 21.0 x 29.7 cm

EUR – Scientific and Technical Research series – ISSN 1018-5593 (print), ISSN 1831-9424 (online)

ISBN 978-92-79-29660-4 (print)
ISBN 978-92-79-29659-8 (pdf)

doi:10.2788/88799

Abstract

Acute oral toxicity is currently being assessed by a suite of refinement test methods based on the traditional LD_{50} lethality test and is, besides skin sensitisation, the only remaining animal test required under REACH Annex VII. In view of assessing the use of alternatives for this health endpoint, EURL ECVAM conducted a study on the 3T3 Neutral Red Uptake cytotoxicity test method addressing the method's capacity to support specifically the identification of substances not requiring classification as acute toxicants. Following independent scientific peer review of this study by EURL ECVAM's scientific advisory committee (ESAC) and having considered input from regulators, stakeholders, international partners and the general public, EURL ECVAM concludes that the 3T3 NRU test method may prove a valuable component of a WoE or ITS approach for supporting hazard identification and safety assessment in agreement with the EU CLP Regulation and international regulatory schemes implementing the upper threshold of UN GHS Category 4 as the cut-off for non-classification of substances. In particular, data from the 3T3 NRU assay may constitute an information source within a WoE approach under the provisions of the REACH regulation (Annex XI, 1.2) potentially supporting conclusions on absence of acute oral toxicity of industrial chemicals.

LB-NA-25946-EN-N

As the Commission's in-house science service, the Joint Research Centre's mission is to provide EU policies with independent, evidence-based scientific and technical support throughout the whole policy cycle.

Working in close cooperation with policy Directorates-General, the JRC addresses key societal challenges while stimulating innovation through developing new standards, methods and tools, and sharing and transferring its know-how to the Member States and international community.

Key policy areas include: environment and climate change; energy and transport; agriculture and food security; health and consumer protection; information society and digital agenda; safety and security including nuclear; all supported through a cross-cutting and multi-disciplinary approach.

Publications Office

ISBN 978-92-79-29659-8

9 789279 296598