OECD Series on Testing and Assessment

Guidance Document on Good In Vitro Method Practices (GIVIMP)

OECD

No. 286

OECD Series on Testing and Assessment

Guidance Document on Good In Vitro Method Practices (GIVIMP)

No. 286

OECD

BETTER POLICIES FOR BETTER LIVES

This document was approved by the Working Group of the National Coordinators of the Test Guidelines Programme in April 2018, and is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology.

This document, as well as any data and any map included herein, are without prejudice to the status of or sovereignty over any territory, to the delimitation of international frontiers and boundaries and to the name of any territory, city or area.

Please cite this publication as:
OECD (2018), *Guidance Document on Good In Vitro Method Practices (GIVIMP)*, OECD Series on Testing and Assessment, No. 286, OECD Publishing, Paris.
https://doi.org/10.1787/9789264304796-en

ISBN 978-92-64-30479-6 (pdf)

Series: OECD Series on Testing and Assessment
ISSN 2077-7876 (online)

The statistical data for Israel are supplied by and under the responsibility of the relevant Israeli authorities. The use of such data by the OECD is without prejudice to the status of the Golan Heights, East Jerusalem and Israeli settlements in the West Bank under the terms of international law.

Photo credits: Cover © science photo/Shutterstock.com

Corrigenda to OECD publications may be found on line at: *www.oecd.org/publishing/corrigenda*.

© OECD 2018

You can copy, download or print OECD content for your own use, and you can include excerpts from OECD publications, databases and multimedia products in your own documents, presentations, blogs, websites and teaching materials, provided that suitable acknowledgement of OECD as source and copyright owner is given. All requests for public or commercial use and translation rights should be submitted to *rights@oecd.org*. Requests for permission to photocopy portions of this material for public or commercial use shall be addressed directly to the Copyright Clearance Center (CCC) at *info@copyright.com* or the Centre français d'exploitation du droit de copie (CFC) at *contact@cfcopies.com*.

About the OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 36 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in twelve different series: Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials; and Adverse Outcome Pathways. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organizations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Foreword

A guidance document on Good *In Vitro* Method Practices (GIVIMP) for the development and implementation of *in vitro* methods for regulatory use in human safety assessment was identified as a high priority requirement by the OECD. The aim of this guidance document is to reduce the uncertainties in cell and tissue-based *in vitro* method derived predictions by applying all necessary good scientific, technical and quality practices from *in vitro* method development to *in vitro* method implementation for regulatory use. This guidance document also applies to *in vitro* methods already accepted by the OECD.

Development of GIVIMP began in 2013 when the OECD Working Group on Good Laboratory Practice (WG GLP) and the Working Group of the National Coordinators of the Test Guidelines Programme (WNT) agreed that there was merit in having guidance for OECD countries on these important issues. The draft guidance was coordinated by the validation body European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM) and was accepted on the work plan of the OECD test guideline programme in April 2015 as a joint activity between the WG GLP and the WNT.

The guidance is targeted primarily at users that implement *in vitro* methods, but also provides guidance for *in vitro* method developers. The document satisfies the following objectives;

1. A detailed update on good practices for state-of-the-art *in vitro* methods applied to regulatory human safety assessment of a variety of compounds.
2. Guidance to users and implementers of *in vitro* methods to help to ensure that Standard Operating Procedures (SOPs) of such methods are well-designed, robust, well-defined and described and can be carried out in a GLP environment, which is essential for use in a regulatory context.
3. Description of the key aspects that may impact the reliability and relevance of the *in vitro* data for quantitative human safety assessment purposes.
4. Description of the importance of reporting criteria, applying good experimental design, establishing acceptance criteria, and performance standards based on scientific evidence from the generated *in vitro* datasets.

The development and revision of GIVIMP has occurred with input from a large group of experts, including experts from both the WG GLP and the WNT. Additionally, an OECD GIVIMP expert group (established specifically for GIVIMP) provided input through teleconferences, a face-to-face meeting and two rounds of written comments.

In January 2017 the first round of comments of the OECD WG GLP and nominated experts of the OECD WNT were forwarded to EURL ECVAM who incorporated these comments, where applicable, and prepared an updated version. The OECD GIVIMP expert group addressed specific outstanding issues on the 23 and 24 March 2017 (Annex D OECD GIVIMP meeting 23-24 March, 2017) and agreed on content, structure, and wording of the GIVIMP so as to provide the OECD with an updated version ready to

enter the second OECD commenting round. After this round, during summer 2017, EURL ECVAM prepared the final GIVIMP version, which was submitted to OECD for adoption.

Table of contents

About the OECD **3**

Foreword **5**

Introduction **13**

Executive summary **17**

Scope **19**

References 21

Chapter 1. Roles and responsibilities **23**

1.1. *In vitro* method developers 24
1.2. Test system providers 25
1.3. Validation bodies 28
1.4. Inter-governmental organisation for cooperation 29
1.5. Suppliers of equipment, materials and reagents 29
1.6. *In vitro* method users 29
1.7. Sponsor 30
1.8. GLP monitoring authorities 31
1.9. Receiving authorities 32
Notes 33
References 34

Chapter 2. Quality considerations **37**

2.1. Quality assurance (QA) and quality control (QC) 38
2.2. Quality risk assessment 39
2.3. Quality control charts 40
2.4. Quality control of test systems 41
2.5. Quality control of consumables and reagents 43
2.6. Staff training and development 44
2.7. Types of documentation 45
2.8. Quality considerations regarding electronic data integrity 45
References 47

Chapter 3. Facilities **49**

3.1. Facility design 50
3.2. Safety, risk assessment and management 52
3.3. Strategies to avoid cross-contamination 55
3.4. Air handling, water supply, environmental control, heating and cooling 57
3.5. Quarantine for new test systems 58
Notes 59
References 60

Chapter 4. Apparatus, material and reagents 61

4.1. Apparatus 62
4.2. Materials and reagents 64
4.3. Use of media in cell culture 65
4.4. The use of antibiotics in cell culture 69
4.5. Additional media components 69
4.6. Dedicated media for particular cell lines 70
Notes 71
References 72

Chapter 5. Test systems 75

5.1. Guidance on Good Cell Culture Practice 76
5.2. Cell and tissue sourcing 77
5.3. Cell and tissue culture transportation 78
5.4. Handling and maintenance of the test system 78
5.5. Cryopreservation and thawing 81
5.6. Cell line identity and genetic aberrations 84
5.7. Contaminants screening: sterility, mycoplasma, virus 86
5.8. Biomarkers and functional tests to confirm the required cell function state 88
5.9. Metabolic activation 89
Notes 91
References 91

Chapter 6. Test and reference/control items 97

6.1. Reference and control items 98
6.2. Applicability and limitations of the method 102
6.3. Test item preparation 104
6.4. Concentration range 104
6.5. Solubility 106
6.6. Stability 109
6.7. Solvents 109
6.8. Air-liquid-interface exposure 110
6.9. Biokinetics/dose extrapolation/interference with media components 110
6.10. Interference with the test system 112
6.11. Interferences in the *in vitro* method 113
6.12. Consideration of interferences not coming from the active ingredient 114
Notes 115
References 116

Chapter 7. Standard operating procedures 121

7.1. *In vitro* method standard operating procedures development 122
7.2. Preparing standard operating procedures 124
References 125

Chapter 8. Performance of the method 127

8.1. Acceptance criteria 128
8.2. Experimental design 129
8.3. In-house validation of the measurement endpoint(s) 134
8.4. Proficiency chemicals 139
8.5. Data-intensive *in vitro* methods 140

Notes ... 141
References ... 142

Chapter 9. Reporting of results ... 145

9.1. Publishing ... 146
9.2. Mutual Acceptance of Data (MAD) ... 147
9.3. Integrated Approaches to Testing and Assessment (IATA) ... 147
9.4. Data reporting for regulatory purposes ... 148
9.5. Reporting of method validation ... 150
Notes ... 151
References ... 152

Chapter 10. Storage and retention of records and materials ... 153

10.1. Data integrity ... 155
10.2. Retention and archiving ... 157
10.3. Backup and restore ... 159
Notes ... 159
References ... 160

Annex A. Good Cell Culture Practice (GCCP) ... 161

Annex B. Good Cell Culture Practice for stem cells and stem-cell-derived models ... 162

Annex C. EURL ECVAM GIVIMP meeting 24-25 February 2015 - participating experts ... 163

Annex D. OECD GIVIMP meeting 23-24 March, 2017 - participating experts ... 164

Annex E. Standardisation and accreditation bodies ... 165

Notes ... 165
References ... 166

Annex F. Good Laboratory Practice ... 167

Notes ... 167
References ... 168

Annex G. Solubility ... 169

Annex H. Biokinetics and xenobiotic bioavailability ... 172

Evaporation / plastic and glass binding / sorption ... 172
Chemical degradation by hydrolyses and phototoxicity ... 173
Metabolism/metabolic stability ... 173
Protein binding ... 173
Cell membrane absorption ... 174
Measurement of free concentration/passive dosing ... 174
References ... 178

Annex I. List of viability testing methods (non-inclusive) of cell cultures ... 181

Glossary ... 187

Notes ... 200
References ... 201

Tables

Table 1.1. Cell culture collections (banks) ... 26
Table 1.2. Examples of data requirements to be documented concerning the origins of cells and tissues ... 27
Table 2.1. Applicability of integrity checks on cell cultures ... 42
Table 3.1. Classification of infective microorganisms by risk group ... 53
Table 3.2. Summary of Biosafety Level requirements ... 53
Table 3.3. Classification of laminar flow biological safety cabinet ... 55
Table 4.1. Serum-free media ... 67
Table 5.1. Comparison of ultra-low cryostorage methods for cells ... 83
Table 5.2. Current status of SNP, STR, and DNA barcode technologies as standard methods for assessing the identity of cell lines from different species ... 85
Table 5.3. Mycoplasma detection methods, their sensitivity, and advantages and disadvantages ... 88
Table 6.1. Common factors affecting solubility ... 106
Table 6.2. Comparison between solubility determination methods ... 108
Table 8.1. Common quality parameters for quantitative and qualitative analytical methods ... 135
Table 8.2. Possible outcomes of an *in vitro* method result of a test item in a validation ... 139
Table 10.1. Terms associated with ALCOA ... 155

Figures

Figure 2.1. QC trend chart for pipette checking ... 40
Figure 3.1. Flow of staff and materials to show separation of processes ... 51
Figure 5.1. Growth curve for cells grown in culture ... 79
Figure 5.2. Flexible two tiered approach to cell banking ... 84
Figure 6.1. BCOP ethanol positive control QC chart March 2002 to November 2005 ... 100
Figure 6.2. Historical data from the Draize Eye Irritation test on two chemicals at 24 hours after instillation where no positive control was included ... 101
Figure 6.3. Neutral red uptake cytotoxicity assay using human keratinocytes ... 102
Figure 6.4. Serial dilutions scheme for KeratinoSens™ test method ... 105
Figure 6.5. Schematic representation of some processes that can cause the final target concentration to be different than the nominal concentration in an *in vitro* test ... 111
Figure 7.1. Evolution of a Standard Operating Procedure (SOP) ... 123
Figure 8.1. Example of plate layout including reference and control items, and solvent and untreated controls ... 131
Figure 8.2. Plate layout for systematic cell seeding errors ... 131
Figure 8.3. Example, Analysis of Competitive Binding Data, with and without use of 10% rule ... 134

Follow OECD Publications on:

http://twitter.com/OECD_Pubs

http://www.facebook.com/OECDPublications

http://www.linkedin.com/groups/OECD-Publications-4645871

http://www.youtube.com/oecdilibrary

http://www.oecd.org/oecddirect/

Introduction

There is a scientific policy and regulatory desire for validated and internationally accepted *in vitro* methods (e.g., OECD test guidelines or ISO standards). To accommodate the needs of receiving authorities, a number of *in vitro* methods, often based on the use of human cells and tissues, have been submitted to international validation bodies during the last two decennia. It was agreed that technical guidance is needed to standardise and advance the development of robust and reliable *in vitro* methods suitable for regulatory purposes. The OECD approached EURL ECVAM to coordinate the drafting of GIVIMP for the development and implementation of *in vitro* methods for regulatory use in human safety assessment which is also equally applicable to non-guideline or not internationally recognised *in vitro* methods.

An Expert Group was established in 2015 to develop such a guidance document. The first draft of the guidance document was prepared following a GIVIMP meeting on the 24 and 25 February 2015 in Ispra, Italy (Annex C: EURL ECVAM GIVIMP meeting 24-25 February 2015) with additional input from experts who could not be present at the meeting. Expert input was received from EURL ECVAM, European receiving authorities (European Food Safety Authority EFSA, European Medicine Agency EMA, European Chemicals Agency ECHA), from the European Union Network of Laboratories for the Validation of Alternative Methods (EU-NETVAL, e.g., from the Belgian, Dutch, Italian, Spanish and Swedish laboratories), from ECVAM's Stakeholder Forum (ESTAF, e.g., the European Society of *In vitro* Toxicology), from the EU and OECD Working Group on Good Laboratory Practice GLP (e.g., delegates from Belgium, The Netherlands, The United Kingdom, Poland, Italy, France, Singapore), from Replacement, Reduction and Refinement (3Rs) Centres (e.g., Centre for Alternatives to Animal Testing, CAAT), from regulatory agencies (e.g., RIVM), from scientists from large industries, Small and Medium Enterprises (SMEs) and from international scientists with expertise in stem cells, cell biology, GLP and *in vitro* methods.

Following the first OECD commenting round of the draft document that took place in the autumn of 2016, a draft version, revised by the OECD Working Group on GLP (WG GLP) and the nominated experts from the Working Group of the National Coordinators of the Test Guidelines Programme (WNT), was circulated in September and November 2016 for review by all 37 members of EU-NETVAL. The OECD GIVIMP Expert Group provided input through teleconferences, a face-to-face meeting (Annex D: OECD GIVIMP meeting 23-24 March, 2017), and two rounds of written comment. The group agreed on content, structure, and wording of the GIVIMP. A further commenting round following revision of Chapter 4.3 on media (specifically, the use of animal serum in cell culture) and Chapter 8 on method performance, by the VMG-NA and GIVIMP Expert Groups, took place in November 2017.

GIVIMP has been updated by Sandra Coecke and Gerard Bowe, and the members of the OECD GIVIMP expert group.

The main authors drafting the version used for the first commenting round are:

Maria Antonietta Antonelli Laura Pioppo	European Medicines Agency (EMA), London, United Kingdom.
Maria Astridou Annamaria Rossi	European Food Safety Authority (EFSA); Parma, Italy.
Bohumila Bichlmaier Suchanová Laura Rossi	European Chemicals Agency (ECHA), Helsinki, Finland.
Camilla Bernasconi Ann-Charlotte Bostroem Gerard Bowe Sandra Coecke Thomas Cole Elisabeth Joossens Ingrid Langezaal Anne Milcamps Sander van der Linden Maurice Whelan	European Commission, Directorate General Joint Research Centre, Directorate F-Health, Consumers and Reference Materials, F.3 Chemicals Safety and alternative Methods, The European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM), Ispra, Italy.
Hui Da Chong	Singapore GLP Monitoring Authority, SPRING Singapore -– Singapore Accreditation Council.
Rita Cortvrindt	Vito, Mol, Belgium.
Laureano Cuevas	Centro Nacional de Microbiologia, Majadahonda, Spain.
Agnieszka Dudra	Bureau for Chemical Substances, Lodz, Poland.
Eugene Elmore	University of California Irvine, Trabuco Canyon, Unites States.
Chantra Eskes	SeCAM, Magliaso, Switzerland.
Kristina Fant Emma Pedersen	RISE Research Institutes of Sweden AB, Borås, Sweden.

Andrew Gray	United Kingdom GLP Monitoring Authority, The Medicines and Healthcare products Regulatory Agency (MHRA), London, United Kingdom.
John Harbell	JHarbell Consulting LCC, Dallas, United States.
Minne Heringa	RIVM, The National Institute for Public Health and the Environment, Centre for Safety of Substances and Products, Bilthoven, The Netherlands.
Delyan Ivanov	University of Nottingham, Cancer Biology Queens Medical Centre, Nottingham, United Kingdom.
Guido Jacobs	WIV-ISP, Scientific Institute of Public Health, GLP-Monitorate-Brussels, Belgium.
Rob Jaspers	Health Care Inspectorate, Ministry of Health, Welfare and Sport, Utrecht, The Netherlands.
Marcel Leist	University of Konstanz, Konstanz, Germany.
Thomas Lucotte	ANSM, Inspectorate Department, Saint-Denis, France.
Richard McFarland	Policy Office of Cellular, Tissue, and Gene Therapies Center for Biologics Evaluation and Research United States Food and Drug Administration, USA
Marisa Meloni	Vitroscreen SRL, Milano, Italy.
Maurits-Jan Prinz	European Commission, Directorate General Internal Market, Industry, Entrepreneurship and SMEs, Brussels, Belgium.
Glyn N Stacey	United Kingdom Stem Cell Bank, Advanced Therapies Division, NIBSC, Medicines and Healthcare products Agency, United Kingdom.
Emanuela Testai	Istituto Superiore di Sanita, Rome, Italy.

Frederique van Acker	Triskelion, Zeist, The Netherlands.
Iwona Wilk-Zasadna	Novartis Vaccines, Basel, Switzerland.

Acknowledgements are also made to the EURL ECVAM colleagues that reviewed and gave useful additional input: Susanne Belz, Mounir Bouhifd, Laura Gribaldo, Tomislav Horvat, Tracey Holley, Annett Janusch Roi, Roman Liska, Alfonso Lostia, Agnieszka Lidia Swiatek, Nicholaos Parissis, Francesca Pistollata, Anna Price, Clemens Wittwehr, Andrew Worth.

Executive summary

The rapid expansion in *in vitro* methods, along with improved understanding of the biological processes involved in toxicological sequelae, have facilitated the development of a variety of predictive *in vitro* methods. These *in vitro* methods can be robust alternatives to using animals to identify and characterise chemical safety hazards. In some regulatory sectors, recent changes in regulation now accept, or in some cases, require, *in vitro* data in lieu of data from animal studies. When the scientific suitability has been demonstrated, the use of *in vitro* methods can reduce the resources required and increase the efficacy of chemical safety evaluation. In order for these alternatives to be used in regulatory decision making the scientific integrity and quality assurance must be assured.

The Guidance Document on Good *In Vitro* Method Practices (GIVIMP) for the development and implementation of *in vitro* methods for regulatory use in human safety assessment was developed as a reference for best practices and as a tool to avoid a reproducibility crisis in *in vitro* toxicological science. The project was a joint activity of the Working Group on Good Laboratory Practises and the Working Group of the National Coordinators to the Test Guidelines Programme. The document includes guidance for developing and using *in vitro* methods for chemical safety assessment, as well as guidance for the laboratory environment in which test data are generated and recorded. The project was coordinated by the European Union Reference Laboratory for alternatives to animal testing (EURL ECVAM) of the European Commission's Joint Research Centre (JRC) with the aim of reducing the uncertainties in cell and tissue-based *in vitro* method derived predictions. The GIVIMP includes a glossary of terms to assure developers and end users begin with a common understanding and tackles ten important aspects related to *in vitro* work.

Chapter 1: Roles and responsibilities, describes the roles of key players over the life cycle of *in vitro* method development and use for safety assessment and provides guidance for improving regulatory acceptance of the method and resulting data. Chapter 2: Quality considerations discusses requirements for development and implementation of *in vitro* methods and considerations to assure the integrity of resulting data. Chapter 3: Facilities, details considerations for the physical environment in which *in vitro* cell and tissue culture are performed to limit impacts that may adversely impact the science. Chapter 4: Apparatus, material and reagents indicates quality requirements for equipment and reagents and includes recommendations to improve reproducibility of the method and results. Chapter 5: Test systems, describes best practices for storage, handling, authentication and characterisation of cell and tissue-based test systems. Chapter 6: Test and reference/control items, provides information on how to assess test system and test item interactions to assure accurate and reliable exposure and avoid *in vitro* method interference due to insolubility and other limitations. Chapter 7: Standard operating procedures, recommends a process for simplifying the work of personnel using the *in vitro* method to assure similar process is followed each time the *in vitro* method is used and reduce variability due to deviations from a fixed methodology. Chapter 8: Performance of the method, describes elements of the experimental design such as plate

layout, data analysis, assessment of linearity, and accuracy to ensure the method is preformed correctly and the endpoint is reliable. Chapter 9: Reporting of results includes guidance on including appropriate detail and recording practises in scientific publications, as well as all related documents, to improve transparency and reproducibility of the method and results. Chapter 10: Storage and retention of records discusses requirements for traceability, storage, verification and transmission of data throughout the life cycle of the *in vitro* method. Key messages and content are highlighted at the beginning of each of the ten sections. Also included in the GIVIMP are eight annexes that provide detailed and directed guidance on specific topics related to good *in vitro* methods practices.

Scope

The major goal of GIVIMP consists of improving the reliability and robustness of *in vitro* methods, reducing the uncertainties of *in vitro* based predictions and therefore increasing the acceptance of the *in vitro* estimated safety measures by regulatory agencies. The scope of the GIVIMP guidance, taking into account good scientific, technical and quality practices, is to ensure that the overall process, starting from *in vitro* method development to the final *in vitro* method implementation for regulatory use is more efficient and effective.

The document emphasis is mainly on human safety assessment using mammalian cell and tissue cultures. It may, however, be broadened to other fields such as environmental safety assessment, gene therapy and immunology domains. It is mainly focused on more commonly used 2D cell and tissue culture systems, but may also be applied to other test systems such as 3D cultures, whole organ systems etc. (Fennema et al., 2013[1]); (Matsusaki, Case and Akashi, 2014[2]).

The document applies mainly to current test systems, practices, trends and processes. If and when felt relevant the WNT will be tasked with issuing a new version. In the various chapters different types of 2D and 3D test systems (cell lines, co-cultures, primary cells, stem cells and tissue cultures) have been provided as examples, however it should also be stated that there are still some reliability issues with the use of some of these test systems for current regulatory testing (e.g., the current "irreproducibility epidemic", challenging scientific questions related to 3D systems) (Frye et al., 2015[3]). Therefore, there was a consensus amongst the OECD GIVIMP expert group that some of the more complex test systems may not yet be at the level required for use in the OECD test guidelines programme, however they may be accepted in the future when the reliability issues are worked out.

In this guidance the OECD Good Laboratory Practice (GLP) term test item is used, where possible, since its applicability ranges from pure substances, mixtures, multi-constituent substances to other types of test items (e.g., nanoparticles, medical devices).

This guidance document targets all players involved in the process, e.g., *in vitro* method developers, *in vitro* test system producers, validation bodies, producers of equipment, materials and reagents, *in vitro* method users, testing laboratories, large industries and small to medium enterprises as well as receiving authorities, monitoring authorities, accreditation bodies and the OECD. The guidance aims to further facilitate the application of the OECD Mutual Acceptance of Data (MAD) agreement to data generated by *in vitro* methods and as such contribute to avoiding unnecessary duplicate testing. This guidance describes the areas related to *in vitro* method development, standardisation, harmonisation, and international acceptance that would benefit from more detailed scientific, technical and quality guidance.

The GIVIMP document has been written with different users in mind, including GLP test facilities but also research laboratories developing new *in vitro* methods. In the latter

case, full compliance with GIVIMP may not be realistic, but compliance with as many as possible of the "good practices" will facilitate the acceptance and routine use of the *in vitro* method in a regulatory environment.

GIVIMP is not intended to duplicate or replace existing OECD guidance or advisory documents but is complementary, addresses specific gaps and aims to collect available references and information on best scientific, technical and quality practices in one document.

GIVIMP is divided into ten sections covering:

1. Roles and Responsibilities
2. Quality considerations
3. Facilities
4. Apparatus, material and reagents
5. Test systems
6. Test and reference/control items
7. Standard operating procedures
8. Performance of the method
9. Reporting of results
10. Storage and retention of records and materials

At the beginning of each chapter a summary box with the key message, key content, guidance for improved practice, and recommendations is included. Abbreviations are repeated per chapter since some readers might read only one chapter, i.e. each chapter may be considered as a separate document. Abbreviations are presented in full at the first occurrence per chapter, For the remainder of the chapter only the abbreviation is used.

Throughout this document, the word *must* is used to denote an obligation; instances of *must* are also often specific to particular context. We use the word *should (be)* to convey a recommendation and that there may exist valid reasons in particular circumstances to disregard the recommendation, but the full implications must be understood and carefully weighed (and documented). The word *may (be)* is generally used to convey an advice and is as such truly optional.

References

Fennema, E. et al. (2013), "Spheroid culture as a tool for creating 3D complex tissues", *Trends in Biotechnology*, Vol. 31/2, pp. 108-115, http://dx.doi.org/10.1016/j.tibtech.2012.12.003. [1]

Frye, S. et al. (2015), "Tackling reproducibility in academic preclinical drug discovery", *Nature Reviews Drug Discovery*, Vol. 14/11, pp. 733-734. [3]

Matsusaki, M., C. Case and M. Akashi (2014), "Three-dimensional cell culture technique and pathophysiology", *Advanced Drug Delivery Reviews*, Vol. 74, pp. 95-103. [2]

Chapter 1. Roles and responsibilities

Key message: *The in vitro method life cycle from development to their use for safety assessment purposes has a variety of key actors and the guidance identifies clearly their responsibilities, both individually and collectively.*

Key content: *Describes all target groups involved in the process e.g., in vitro method developers, test system (cells, tissues) providers, validation bodies, inter-governmental organisations for cooperation, suppliers of equipment, materials and reagents, in vitro method users (e.g., testing laboratories, large industries and small to medium enterprises), sponsors, receiving authorities and GLP monitoring authorities.*

Guidance for improved practice: *Besides the elements necessary for good scientific work in discovery, additional requirements, such as documentation, ownership, identity and genetic make-up, related to the in vitro method and the test system, are key for their in vitro method acceptance at regulatory level.*

Recommendations *are given for several of the target groups on how to put into practice their responsibilities for facilitating the development and implementation of in vitro methods for regulatory use.*

In vitro methods are often developed without the primary aim of being used for regulatory purposes, but are rather focused on the discovery of disease pathways or investigation of mechanisms of action induced by external factors causing cell disturbance. These methods however, can form the basis for future *in vitro* methods used either for safety assessment or for toxicity screening.

Many of the following organisations (e.g. validation bodies, receiving authorities) should not be considered as single entities but consist of a network of advisory (e.g. scientific, technical, ethical) bodies which feeds into the processes and roles detailed below.

1.1. *In vitro* method developers

In vitro method developer(s) refers to the person or entity who develops or has developed an *in vitro* method destined for regulatory use in human safety assessment.

Researchers aiming to develop *in vitro* methods suitable for regulatory testing purposes must be aware that beyond the definition, description and within-laboratory repeatability and reproducibility of the *in vitro* method, receiving authorities have additional quality requirements for test acceptance (OECD, 2005[1]). The *in vitro* method developer should keep in mind that the quality of historical data and documentation regarding the *in vitro* method will have a significant impact on the regulatory acceptance process.

Briefly, the *in vitro* method developer is responsible for providing a clearly written and well documented *in vitro* method description, and related Standard Operating Procedure(s) (SOP(s)), taking into consideration all aspects described in GIVIMP.

The developer's knowledge and understanding of the *in vitro* method is the basis for establishing an approach to control the *in vitro* method and to set for instance adequate acceptance criteria for the results obtained when running the *in vitro* method. Each developer should judge whether he or she has gained sufficient understanding of the *in vitro* method to provide a high degree of assurance to successfully propose the *in vitro* method for regulatory applications.

In vitro method developers should take into account the Intellectual Property (IP) guidelines and good licensing practices regarding test systems as set out on the OECD website[1] [2]. The Guidelines for the Licensing of Genetic Inventions[3] were adopted by OECD member countries in 2006. Although the Guidelines describe the principles and best practices for the licensing of genetic inventions used in human health care, the principles can generally be promoted in other areas in the field of regulatory testing of chemicals for the protection of human health and the environment. Currently an OECD guidance on best practices for licensing of protected elements in OECD test guidelines is in development.

New test guidelines proposals should provide information on Intellectual Property Rights (IPR) aspects, as transparently as possible. In particular, the following information is expected to be provided: "Describe if the *in vitro* method includes components, equipment or other scientific procedures that are covered (or pending) by IPR (e.g., patents, patent applications, industrial designs and trademarks) and/or intended to remain confidential. Information should be provided on the overall availability of the IPR-protected components including whether they are commercially available or require a Material Transfer Agreement or other licensing agreements. In addition, the possibility of providing a generic description of the IPR-covered component/test system as well as any

other element intended to remain confidential should be disclosed and whether Performance Standards have been developed for the *in vitro* method."[4]

The development of Performance Standards[5] was agreed as the solution to overcome concerns regarding market monopoly (e.g. where a commercial provider could take a disproportionate financial advantage due to the inclusion of proprietary elements in test guidelines). The development of Performance Standards will also enable the development of similar test methods and facilitate their validation.

As the use of mammalian, including human, cells and tissues is critical for the development and implementation of *in vitro* methods for regulatory use in human safety assessment, already in the early stages care has to be taken regarding their ownership, their identity and genetic make-up and who can control their fate. A number of treaties, laws, and regulations help to guide the ethical collection of human-derived specimens (Allen et al., 2010[2]).

Reference data to assess the relevance of *in vitro* methods are typically from surrogate animal studies ("*in vivo* animal data"), but can also be derived from other sources. This is especially important for areas where the mechanism of action is not preserved across species, e.g. metabolism, CYP induction (Sinz, Wallace and Sahi, 2008[3]), and where the availability of human reference data for the mechanism studied is essential. Human data can be obtained from epidemiological, clinical or other resources. In the case of prospective generation of human reference data, approval will need to be sought from an independent committee subject to national laws[6].

When *in vitro* method developers conclude their *in vitro* method is sufficiently developed, they can then proceed to an in-house performance assessment (see Section 8.3). When such internal assessment is successful, they can submit the *in vitro* method to a validation body for the formal validation of the method, or, can organise the validation by themselves. The *in vitro* method developers should be able to prove that the *in vitro* method they offer to the validation body is robust, reliable, relevant, and supported by high quality data as described in the present guidance.

In order to have the *in vitro* method considered for regulatory acceptance, the *in vitro* method developer needs to contact the appropriate national coordinator[7] to prepare a Standard Project Submission Form (SPSF) for a new Test Guideline, or in the case of 'me-too' *in vitro* methods for addition to the relevant Performance Based Test Guideline (PBTG) (OECD, 2016[4]). Project proposals for new Test Guidelines need the active support of receiving authorities in at least one member country, and have to meet a regulatory need in member countries.

1.2. Test system providers

In vitro test systems are mainly biological systems, quite often consisting of tissues or cell lines. Test systems can be developed in-house (i.e. by the *in vitro* method developer), acquired from other laboratories or purchased from a cell culture bank, either academic or commercial. The OECD consensus document, Compliance of Laboratory Suppliers with GLP Principles, recommends that test system providers should adhere to a formal quality system, such as International Standard ISO 9001 (OECD, 2000[5]).

The responsibility for the quality and documentation of the test system rests entirely with the test facility (Section 5.2), however, the role of the supplier is crucial in aiding the facility meet these quality requirements, e.g. test systems characterisation requirements

can often be directly fulfilled by information from the supplier (OECD, 2000[5]). Accredited/certified providers generally provide extensive documentation on the origins and characterisation of the test system and may also offer advice/services, such as quality assurance guidance, cell culture maintenance, and safety practices for use and disposal of the test system, including transport and containment[8] (OECD, 2004[6]) (Coecke et al., 2005[7]).

It is difficult, if not impossible, to identify cell lines from different origins and ensure that they are not cross-contaminated, misidentified or mixed-up (The European Collection of Authenticated Cell Cultures ECACC Handbook – Fundamental Techniques for ECACC Cell Lines[8]), based solely on morphology and/or culture characteristics (Section 5.6). Infection or contamination of a cell line with an adventitious virus or mycoplasma may significantly change the characteristics of the cells but again such contamination may not be visibly evident. The test system provider should therefore provide documentation the cell line's authenticity including verification of its identity and proof to be free of cross-contamination by other cell lines (Section 5.6) and/or contamination caused by bacteria, yeast or fungi, mycoplasma (Section 5.7). Additional information on the origin and culture history of the cell line, ideally including its transfer among laboratories and repositories, its manipulation (physicochemical or genetic), and details on the types of tests carried out for the detection and (if applicable) elimination of contamination should be made available, so as to provide complete tracking of the cell line provenance. In some cases, e.g., cell lines established many years ago may lack some aspects of their provenance and their origin may be unknown. It is therefore recommended to confirm that the cells in current use are assessed against a previously authenticated stock (where available), either in a cell bank or in the laboratory of the originator.

Test systems sourced from recognised cell culture banks (Table 1.1) are unlikely to be contaminated with microorganisms, unless stated otherwise, and generally provide adequate documentation, usually in the form of a Certificate of Analysis, including a Short Tandem Repeat (STR) profile.

Table 1.1. Cell culture collections (banks)

Cell culture collections	Country	Web site
American Type Culture Collection (ATCC)	USA	http://www.atcc.org
CellBank Australia	Australia	www.cellbankaustralia.com
Coriell Cell Repository	USA	http://locus.umdnj.edu/ccr
Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSMZ)	Germany	http://www.dsmz.de
European Collection of Animal Cell Cultures (ECACC)	UK	http://www.camr.org.uk
Japanese Collection of Research Bioresources (JCRB)	Japan	http://cellbank.nihs.go.jp
RIKEN Gene Bank	Japan	http://www.rtc.riken.go.jp
UK Stem Cell Bank (UKSCB)	UK	http://www.nibsc.org/ukstemcellbank

The test system provider must provide all relevant safety information, in compliance with national and international regulations, for the transport, use and disposal, including containment in the case of an accident. Where the *in vitro* method developer is also the test system provider or where the test system has been acquired from other laboratories, the *in vitro* method developer must ensure the availability of the test system both in the short and long term and as such take on all responsibilities associated with a test system provider regarding documentation and quality control.

The Guidance on Good Cell Culture Practice: A Report of the Second ECVAM Task Force on Good Cell Culture Practice (GCCP) (Coecke et al., 2005[7])provides a minimal set of requirements for documentation. However the documentation requirements listed in Table 1.2 may not be feasible in all cases when working with cells or tissues of animal or human origin, and in particular when animal-derived tissues are obtained from abattoir operations[9]. The OECD GLP document No 14 (The Application of the Principles of GLP to *in vitro* Studies) states that some characteristics of the test systems can be fulfilled with assistance from the supplier, however the performance when evaluated with appropriate reference items, including positive, negative, and untreated and/or vehicle controls, where necessary, is the responsibility of the relevant study director (OECD, 2004[6]).

Table 1.2. Examples of data requirements to be documented concerning the origins of cells and tissues

	Isolated organs and tissues of animal origin	Primary cultures of animal origin	All materials of human origin	Cell lines
Ethical and safety issues	+	+	+	Applicable, if human or involving recombinant DNA or pathogens
Species/strain	+	+	+	+
Source	+	+	+	+
Sex	+	+	+	+
Age	+	+	+	+
Number of donors	+	+	If applicable	na
Health status	+	+	+	+
Any special pre-treatment	+	+	+	+
Organ/tissue of origin	+	+	+	+
Cell type(s) isolated	+	+	+	+
Isolation technique	+	+	+	+
Date of isolation	+	+	+	+
Operator	+	+	+	+
Supplier	+	+	+	+
Informed consent	na	na	+	If human, may be applicable
Material transfer agreement	na	na	+	+
Medical history of donor	na	na	+ (if available)	If human, may be applicable (if available)
Pathogen testing	If applicable [a]	If applicable [a]	+[a]	+[a]
Shipping conditions	+	+	+	+
State of material on arrival	+	+	+	+
Biosafety classification	+	+	+	+
Cell line identification and authentication	na	na	na	+
Mycoplasma testing	na	na [b]	na [b]	+

Notes:
1. Screening tests for animal colonies or donors of cells and tissue may be appropriate.
2. May be important if material is preserved for longer term use (e.g. as feeder layers for other cultures).

na = not applicable
Source: (Coecke et al., 2005[7])

1.3. Validation bodies

The role of national and international organisations, such as OECD related working groups, EURL ECVAM, ICCVAM, JaCVAM, Health Canada, KoCVAM, etc., is to promote and facilitate *in vitro* method validation for regulatory acceptance to replace, reduce or refine (3Rs) *in vivo* testing. The validation body's responsibility is to contribute to both an effective validation process and to ensure the quality of the validated *in vitro* method.

The basic principle of validation is to assess that an *in vitro* method is fit for its intended use. To this end, the validation process generally consists of the generation, collection and evaluation of data to establish scientific evidence that the *in vitro* method is capable of consistently producing data that is reliable (reproducible) and relevant for the intended purpose. For further information regarding validation concepts, challenges, processes and useful tools see Chapter 04 in Validation of Alternative Methods for Toxicity Testing (Griesinger et al., 2016[8]).

While details can differ between validation bodies, the overall goal of the process of validation is to improve the international acceptance of test methods. To this end, the OECD has drafted a guidance document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment No. 34 (OECD, 2005[1]). The document promotes a "modular approach to validation", where the information needed to support the validity of a test method is organised into modules (Hartung, 2004[9]). Several practical aspects need to be considered in the design and validation of *in vitro* methods (Coecke et al., 2016[10]), if the ultimate aim is to generate a dataset that can support the development of an international test guideline. OECD GD 34 provides information on the following aspects:

- test definition (including purpose, need and scientific basis);
- within-laboratory repeatability and reproducibility;
- between-laboratory transferability;
- between-laboratory reproducibility;
- predictive capacity (accuracy);
- applicability domain; and,
- performance standards.

Although validation is an important step, not all modules/aspects of validation are indispensable for regulatory acceptance. It is important to emphasise that only robust methods can be accepted, i.e. reproducibility and transferability have to be demonstrated, thus validation is not entirely indispensable. After successfully demonstrating the validity of an *in vitro* method to a validation body, the method can be presented to the OECD for regulatory acceptance, depending on the Member State, e.g. in the US, it must be posted in the Federal Register for comment and then FDA and EPA need to separately evaluate comments and follow up by posting final guidance in the federal register. Once *in vitro* methods are consolidated within an OECD test guideline (TG), data produced by using those methods are mutually accepted by all OECD Members and MAD-adhering Country Authorities, unless specific national regulatory requirements are not met.

1.4. Inter-governmental organisation for cooperation

A framework for cooperation between inter-governmental organisation was established in the critical areas of validation studies, independent peer review, and development of harmonised recommendations to ensure that alternative methods/strategies are more readily accepted worldwide.

The International Cooperation on Alternative Test Methods (ICATM)[10,11,12] was formally established in 2009 through a collaboration involving EU, US, ,Japan and Canada. Representatives now include, EU (EURL ECVAM), the US (NICEATM/ICCVAM), Japan (JaCVAM), Canada (Health Canada), and Korea (KoCVAM). Although not yet formally partners, the Brazilian Center for the Validation of Alternative Methods (BraCVAM) and China also actively participate.

ICATM partners are working together to promote enhanced international cooperation and coordination on the scientific development, validation and regulatory use of alternative approaches.

1.5. Suppliers of equipment, materials and reagents

While the responsibility for the quality and fitness for use of equipment and materials rests entirely with the management of the test facility, it is in the suppliers' interests to meet these requirements where possible. Suppliers are recommended to comply with formal national or international standards or to be accredited within various national schemes, where appropriate (OECD, 2000[5]). Selection of suppliers should follow a formal documented process and should be reviewed regularly to ensure that equipment, materials and reagents meet the facility's requirements.

When performing established *in vitro* testing methods, the test results can only be accepted if the equipment, materials and reagents used are of proven quality as established by formal testing or evaluation procedures.

Equipment suppliers should provide all information necessary to operate and maintain the equipment, including equipment and software manuals and quality and safety conformation certificates and warranties. For complex equipment it is recommended that the manufacturer install the equipment and provide the necessary documentation to confirm the correct functioning of the equipment according to the manufacturer's specifications (Section 4.1).

Characteristics of the supplied materials and reagents should be appropriately documented in adequate quality documents such as a certificate of analysis, batch release certificate or similar.

1.6. *In vitro* method users

In vitro method user(s) herein refers to the person(s) or entity that uses the finalised *in vitro* method. As the final goal of these *in vitro* methods is to be included in a regulatory framework, the majority of these users will be GLP compliant test facilities. GLP test facilities' responsibilities are described in the OECD Principles of GLP[13] or equivalent GLP principles as defined in national legislation.

In vitro method users should document their competency to perform a test in compliance with a specific OECD TG, e.g., by running the proficiency chemicals (Section 8.4.) or checking the performance of the method (Chapter 8).

Non-GLP *in vitro* method users can also profit from the use of the GIVIMP guidance. In these cases no regulations exist and no responsibilities are defined, however it is highly recommended to apply all necessary good scientific, technical and quality practices that this guidance describes so as to reduce the uncertainties in the results produced by the *in vitro* method. Appropriate accreditation, e.g., ISO/IEC 17025[14], may be requested or recommended in some cases.

The responsibility for the quality, integrity and compliance (where applicable) of all data generated and reported rests entirely with the *in vitro* method user(s), who must also verify and assure that the quality of all products and materials used in the generation of said data meets the required specifications as described in the *in vitro* method and/or other regulatory guidelines. To be able to prove this, *in vitro* method users will need to work with preferred or approved suppliers who are selected on predefined criteria (e.g., ISO certification, controlled transport, technical support, assured delivery, batch selection, etc.).

1.7. Sponsor

Studies are often initiated by a sponsor who is responsible for ensuring that a study is conducted according to certain requirements e.g., GLP (OECD, 1998[11]).

The sponsor should actively verify that the study is conducted in accordance with all Principles of GLP. The sponsor should verify that the involved test facility including, if applicable, any test sites are able to conduct the study in accordance with the GLP Principles. For example, the sponsor could monitor the involved test facility and, if applicable, any other test site also involved in the study, prior to and during the conduct of the study. In addition, the sponsor might also check the compliance status of a test facility as determined by the national GLP compliance monitoring authority.

The sponsor should be aware, however, that only the study director remains ultimately responsible for the scientific validity and the GLP compliance of the study.

The sponsor should also ensure the integrity of each unaltered study report submitted to receiving authorities.

- The sponsor may be responsible for providing the test item. To ensure that there is no mix-up of test items the sponsor should, in cooperation with the test facility, define a mechanism to allow verification of the identity of the test item for each study.
- Often the sponsor is responsible for characterisation of the test item. In that case, the study director should ensure that this is explicitly mentioned in the study report.
- Where the sponsor is responsible for the characterisation of the test item, the sponsor is expected to disclose all information regarding the characterisation of the test item to the study director, and should be explicitly stated in both the study plan and the final report.

- The sponsor should inform the test facility about any potential risks of the test item to human health and environment as well as any necessary protective measures and disposal procedures.
- In some countries the sponsor should formally approve the study plan by dated signature.
- The name and full address of the sponsor should be mentioned in the study plan and study report.
- Where the study materials including study plan, raw data, specimens and samples of test and reference items and final reports are transferred to the sponsor, the sponsor assumes the responsible for ensuring that all materials are archived in accordance with the GLP Principles.

On the basis of the outcome of the studies the sponsor may decide to submit a test item for registration to the receiving authorities.

1.8. GLP monitoring authorities

GLP Monitoring Authorities (MAs) are established by the governments of OECD Member States and MAD-adherent countries. Some countries have only one MA, while others have more than one e.g., in Japan there are eight MAs while in the US there are two MAs, the US Environmental Protection Agency (EPA) and the US Food and Drug Administration (FDA). The OECD maintains a list of links to national web sites on GLP, including information on MAs[15].

For studies conducted for regulatory purposes, the responsibility for evaluating the results of the study lies with the regulatory reviewer at the receiving authority. However, this evaluation can only be effective if the study data can be relied upon. GLP ensures that the quality and integrity of the data can be demonstrated and the conduct of the study reconstructed.

The OECD expects member countries to establish national MAs, a body or bodies responsible for monitoring the GLP compliance of test facilities within their territories and according to national legal and administrative policies. In the European Union (EU), facilities included in the GLP monitoring programme of the GLP Monitoring Authority are inspected on a regular basis, approximately every two to three years. Routine monitoring inspections also include study audits. In addition, MAs can be requested by a receiving authority to conduct specific study audits as a result of concerns raised following the review of a regulatory submission. The MA has ultimate responsibility for determining the GLP compliance status of test facilities and/or GLP studies. The MA also has responsibility for taking any action based on the results of test facility inspections or study audits which are deemed necessary.

The respective national compliance MAs are also responsible for the exchange of information on the compliance of test facilities inspected, and should provide relevant information concerning their procedures for monitoring compliance. They have the responsibility to facilitate the MAD (Section 9.2) multilateral agreement, which states that test data generated in OECD countries and full adherent countries – (Argentina[16], Brazil, India, Malaysia, South Africa and Singapore)[17] in accordance with OECD Test Guidelines and the OECD Principles of GLP shall be accepted in other member countries

by regulatory bodies for assessment purposes and other uses relating to the protection of human health and the environment[18].

1.9. Receiving authorities

Receiving authorities receive non-clinical safety data as part of regulatory submissions and they must ensure that the legal requirements are met. Receiving authorities include the European Chemicals Agency (ECHA), European Medicines Agency (EMA), European Food Safety Authority (EFSA), as well as various national agencies that are responsible for assessing safety data such as, for example, the US EPA, FDA and Department of Agriculture (USDA) and in Japan the Ministry of Health, Labour and Welfare (MHLW) and the Ministry of Agriculture, Forestry and Fisheries (MAFF).

The responsibility of the receiving authorities is to check that test data are obtained according to available OECD TGs (where applicable) and guidance documents and that they use the data accordingly in their evaluations and according to the regulatory framework. With regard to GLP studies, the receiving authorities verify whether the reported study was conducted in compliance with GLP (Section 1.8). The level of GLP compliance verification depends on the particular receiving authority and the specific legal framework. Receiving authorities may request a study audit if a concern about the GLP compliance status of the study is identified or in case the responsible test facility has not been inspected by the responsible national GLP monitoring authority. Receiving authorities may additionally indicate to *in vitro* method developers where they see a need for new or better methods, and to validation bodies which methods deserve priority in the validation.

In vitro methods are becoming more and more accepted for regulatory use and some regulation requiring toxicological data, allow or even encourage the use of alternative methods. Multiple legislational frameworks, e.g., US Federal agencies (Schechtman, 2002[12]) and EU Directive 2010/63/EU[19], in various regions of the world have statements that include reference to the "3Rs" or that express support for the replacement, reduction, and refinements of animals use where feasible.

The U.S. EPA's Office of Pesticide Programs (OPP) is developing and evaluating alternative approaches to replace or amend more traditional methods of toxicity testing and uses so-called Integrated Approaches to Testing and Assessment (IATA) (see Strategic Vision for Adopting 21st Century Science Methodologies[20]), with the immediate goal to significantly reduce the use of animals in acute effects testing.

As a result of these developments European and national regulatory bodies tend to increasingly accept data generated by alternative methods, especially from validated *in vitro* methods. Data generated using non-validated *in vitro* methods may be accepted as supportive information or when mechanistic data are required. Although the applicability of *in vitro* methods to meet regulatory needs may be different in individual OECD member countries, many countries have adopted the principles of Replace, Reduce and Refine (3Rs) and are proactively supporting the use and implementation of alternative methods[21].

EMA expresses in a number of documents their vision and action plans towards the implementation of the 3Rs principles (EMA, 2014[13]). Besides established formal validation processes by recognised institutions such as the Centres for the Validation of Alternative Methods (CVAMs) and The European Directorate for the Quality of

Medicines & HealthCare (EDQM), multiple and flexible approaches are considered acceptable to demonstrate scientific validity of new testing approaches and their fitness for regulatory use, either as pivotal, supportive or as exploratory mechanistic studies (EMA, 2016[14])

Notes

1. See: http://www.oecd.org/chemicalsafety/testing/intellectual-property-in-oecd-test-guidelines.htm
2. See: http://www.oecd.org/sti/sci-tech/intellectualpropertyinbiotechnologyandthelifesciences.htm
3. See: http://www.oecd.org/science/biotech/36198812.pdf
4. See: http://www.oecd.org/chemicalsafety/testing/intellectual-property-in-oecd-test-guidelines.htm
5. See: http://www.oecd.org/chemicalsafety/testing/performance-standards.htm
6. See: https://www.moh.gov.sg/content/dam/moh_web/Publications/Guidelines/Human%20Biomedical%20Research/2007/IRB%20Operational%20Guidelines_14-12-07_formatted.pdf
7. See: https://www.oecd.org/chemicalsafety/testing/national-coordinators-test-guidelines-programme.htm
8. See: http://www.sigmaaldrich.com/life-science/cell-culture/learning-center/ecacc-handbook.html
9. Justification should be provided when documentation requirements listed in Table 2 are not followed.
10. See: https://ntp.niehs.nih.gov/pubhealth/evalatm/iccvam/international-partnerships/index.html
11. See: https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing/advisory-bodies/icatm
12. See: http://www.jacvam.jp/en_effort/icatm.html
13. See: http://www.oecd.org/chemicalsafety/testing/oecdseriesonprinciplesofgoodlaboratorypracticeglpandcompliancemonitoring.htm
14. See: https://www.iso.org/standard/66912.html
15. See: http://www.oecd.org/chemicalsafety/testing/linkstonationalwebsitesongoodlaboratorypractice.htm
16. See: Full adherence for Argentina only applies to industrial chemicals, pesticides and biocides
17. See: http://www.oecd.org/env/ehs/non-memberadherentstotheoecdsystemformutualacceptanceofchemicalsafetydata.htm
18. See: http://www.oecd.org/env/ehs/mutualacceptanceofdatamad.htm
19. See: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:276:0033:0079:EN:PDF

20. See: https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/strategic-vision-adopting-21st-century-science

21. See: https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/in-vitro-methods

References

Allen, M. et al. (2010), "Human Tissue Ownership and Use in Research: What Laboratorians and Researchers Should Know", *Clinical Chemistry*, Vol. 56/11, pp. 1675-1682, http://dx.doi.org/10.1373/clinchem.2010.150672. [2]

Coecke, S. et al. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33, no. 3, pp. 261-287. [7]

Coecke, S. et al. (2016), "Practical Aspects of Designing and Conducting Validation Studies Involving Multi-study Trials", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_5. [10]

EMA (2016), *Guideline on the principles of regulatory acceptance of 3Rs (replacement, reduction, refinement) testing approaches*. [14]

EMA (2014), *Concept paper on review and update of EMA guidelines to implement best practice with regard to 3Rs (replacement, reduction and refinement) in regulatory testing of medicinal products*. [13]

Griesinger, C. et al. (2016), "Validation of Alternative In Vitro Methods to Animal Testing: Concepts, Challenges, Processes and Tools", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_4. [8]

Hartung, T. (2004), *A modular approach to the ECVAM principles on test validity*, ATLA Alternatives to Laboratory Animals. [9]

OECD (2016), *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264265295-en. [4]

OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, OECD Series on Testing and Assessment, No. 34, OECD Publishing, Paris. [1]

OECD (2004), *The Application of the Principles of GLP to in vitro Studies*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 14, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084971-en. [6]

OECD (2000), *Compliance of Laboratory Suppliers with GLP Principles*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 5, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078611-en. [5]

OECD (1998), *The Role and Responsibility of the Sponsor in the Application of the Principles of GLP*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 11, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078734-en. [11]

Schechtman, L. (2002), "Implementation of the 3Rs (Refinement, Reduction, and Replacement): Validation and Regulatory Acceptance Considerations for Alternative Toxicological Test Methods", *ILAR Journal*, Vol. 43/Suppl_1, pp. S85-S94, http://dx.doi.org/10.1093/ilar.43.suppl_1.s85. [12]

Sinz, M., G. Wallace and J. Sahi (2008), "Current Industrial Practices in Assessing CYP450 Enzyme Induction: Preclinical and Clinical", *The AAPS Journal*, Vol. 10/2, pp. 391-400, http://dx.doi.org/10.1208/s12248-008-9037-4. [3]

Chapter 2. Quality considerations

Key message: *To realise fully the potential of in vitro methods and allowing them to become a key tool for a new way of doing toxicology, they need to be developed and applied in a way that scientific integrity and quality is assured.*

Key content: *Discusses quality assurance versus quality control, quality risk-based assessment and quality control requirements for development and implementation of in vitro methods, the types of documentation needed and quality considerations regarding the integrity of the data.*

Guidance for improved practice: *Control charts can be used as a powerful and simple statistical tool to help routinely monitor the quality of any quantitative process and to determine if the process is in a state of control.*

Recommendations *for basic quality risk assessment questions and applicability of integrity checks on cell and tissue cultures are described.*

The life cycle of an *in vitro* method usually progresses from method development to validation to routine use. *In vitro* method development may benefit from many quality considerations addressed in the GIVIMP guidance document, e.g., recommendations concerning the test system, maintenance and calibration of equipment, qualification of computerised systems and training requirements. *In vitro* method developers, who do not work in a formal quality system, may also benefit from certain Quality Assurance (QA) requirements such as consistent documentation, an internal QA program and change control policies (OECD, 2016[1]). In summary, it is recommended that method validation be performed in a formal quality environment, while routine use of *in vitro* methods for safety testing should always be performed in a formal quality system environment, often meeting the requirements of GLP or similar quality systems.

The International Organization for Standardization (ISO) publishes many standards of which ISO/IEC 17025 is the main standard used by testing and calibration laboratories. ISO/IEC 17025 (General requirements for the competence of testing and calibration laboratories), originated in the laboratory accreditation community who prepared a mutually agreed set of criteria that a laboratory should fulfil in order to demonstrate its technical competence. ISO/IEC 17025 is an international standard that laboratories can choose to apply (i.e. voluntary). Increasingly governments are specifying international standards, such as ISO/IEC 17025, as a tool to meet their regulatory and trade objectives across a wide range of fields (OECD, 2016[2]).

The principles of GLP on the other hand are written into law in many countries as a regulatory control mechanism, often as a legal requirement that non-clinical health and environmental safety studies intended for regulatory submission be conducted under GLP. The OECD GLP Principles have gained wide acceptance, also in non-OECD countries. In 2004 the OECD published an Advisory Document on The Application of the Principles of GLP to *in vitro* Studies, so as to provide guidance specifically of relevance to the application and interpretation of the OECD Principles of GLP to *in vitro* studies.

Even though there is overlap in many areas between GLP and ISO/IEC 17025 (e.g., training, management of equipment, etc.) each serve, as a result of their evolution and history, very different purposes. The OECD Principles of GLP are specifically designed to be applied to individual studies and to accommodate the complexity and variability of non-clinical health and environmental safety studies, while ISO/IEC 17025 was originally intended for testing according to established or specifically developed methodology. However, laboratory accreditation such as ISO/IEC 17025 can be applied to non-clinical testing, and is increasingly being used by governments to meet regulatory and trade objectives.

2.1. Quality assurance (QA) and quality control (QC)

The definition and roles of both QA and QC will depend for a large part on what quality management system is being followed; however most systems have a defined Quality Assurance Unit (QAU) that acts in an independent role. For the sake of simplification QA may be described as a proactive process for managing quality, while QC may be thought of as a reactive process for recognising quality problems and correcting them. The quality management system should be under ongoing review to ensure current best practices are implemented and to provide continuous improvements in the quality system, even if not formally required for GLP.

GLP has no explicit requirement to undertake QC activities, and QC is not defined or included in the GLP Principles or any of OECD GLP consensus or advisory documents. The OECD GLP Principles refer to a Quality Assurance Programme as a defined system, carried out by individual(s) designated and directly responsible to management who must not be involved in the conduct of the study, that is designed to assure that studies performed are in compliance with the principles of GLP.

Most GLP facilities do include QC activities within their quality system. QC activities are most effective when built into a procedure, e.g., calibration or checking of an instruments performance prior to use in order to identify and correct errors at the earliest opportunity prior to acquisition of study data.

2.2. Quality risk assessment

Risk management includes elements such as risk identification, assessment, mitigation, elimination and communication and may be applied to many laboratory processes, such as setting the calibration interval for specific equipment (e.g., some equipment may require less frequent calibration that others based on the probability of failure, the ease of detection and the severity of the consequences of the failure). Quality risk assessment may also be used for the assessment and evaluation of suppliers or to ensure that the test systems (Section 2.4), reagents and materials (Section 2.5) etc. are fit for purpose.

In a risk assessment the following basic questions should be addressed:

- What might go wrong?
- What is the nature of possible risks?
- What is the probability of their occurrence and how easy is it to detect them?
- What are the consequences (the severity)?

For an effective quality risk assessment the probability that the event will occur and the severity of the event must be addressed. Other parameters, such as assessing the ease of detection and the frequency of occurrence, may also be included to provide a more fine-tuned approach. The probability can be based on historical data and/or on the users' experience or it may also remain unknown. The severity of the event is addressed by listing the possible consequences of the event in the case it actually occurs. The ease of detection is a more difficult concept and is usually based on experience and thorough knowledge of the process while the frequency of occurrence may be based on historical data or also remains unknown.

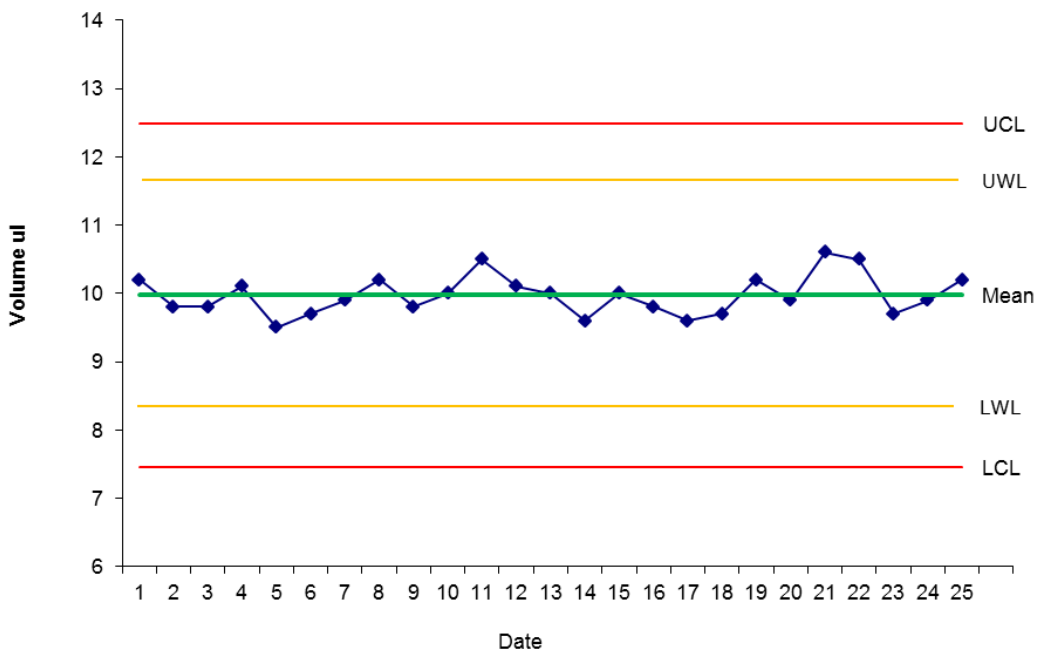
The output of a risk assessment is either a quantitative estimate of risk (numeric probability) or a qualitative description of a range of risk (e.g., high/medium/low). The use of historical data is important when evaluating the probability that the event will occur. Therefore, in order to use updated information, the risk must be reassessed periodically.

Based on the outcome of the analysis and the criticality of the level of risk, specific risk controls, such as increased quality controls or QA inspections, should be put in place. The purpose of these quality risk controls is to reduce the risk to an acceptable level, and should be proportional to the significance of the risk.

2.3. Quality control charts

Control charts may be used in certain QC activities, and are a powerful and simple statistical tool used to routinely monitor the quality of any quantitative process to determine if the process is in a state of control. Control charts are typically used for time-series data, e.g. Figure 2.1 but they may also be used for monitoring discrete data sets such as batch to batch variability or operator performance.

Figure 2.1. QC trend chart for pipette checking



Run or trend charts (Figure 2.1) are the most commonly used and easily understood charts. Individual results, e.g., for a reference item or for pipette checking, are plotted using a scatter plot graph versus the time order when the data were produced. The data points are linked by lines to help visualise the trend or changes in the trend. Trend charts are ideal for visually checking the historical performance of a process. Additional information may be placed on the trend charts to aid decision making, such as the true or expected value and specification limits or the average and control limits may be calculated based on historical data.

When using historical data, which gives a true representation of the performance of the process, it is important that the data used is representative of the current process and is based on an adequate sample size, i.e. the smaller the set of historical data used to calculate the average and limits the less representative these are of the overall process. The average (often used as a running average) is plotted with limits set at ± 2 Standard Deviations (SDs) for Upper and Lower Warning Limits (UWL and LWL) and ± 3 SD for Upper and Lower Control Limits (UCL and LCL). For normally distributed data based on a representative data set, the warning limits contain approximately 95% of the data points while 99.7% of the data are contained within the control limits, i.e. the 68–95–99.7 rule. The limits are irrespective of the process specifications or requirements.

Control or Shewhart charts use subgroups of the individual data to smooth out effects of individual data points and as such make it easier to identify trends or changes in the process. The most common types are the X-chart (average), the R-chart (range) and the s-chart (standard deviation). Subgrouping of the data, e.g., into sets of 5, allow the calculation of the standard deviation and/or range providing more information and finer control of the process. The subgroup average is usually set as the central line and the limits are calculated based on 3 SDs.

Control charts are mainly used to identify when a process is out of control or about to go out of control. When the process is out of control, data points fall outside the control limits, or when the process is about to go out of control, i.e. when a trend (e.g., two consecutive points outside the warning limits but still within the control limits) it usually means a new source of variation has been introduced into the process.

This variation may be due to systematic error which is usually seen as a change in the mean of the control values. Systematic error may be due to an abrupt change in the process (out of control), often caused by a sudden failure (e.g., apparatus), due to operator error (e.g., pipetting error) or some other once-off event. Systematic error may also be due to a gradual change which does not cause the process to go out of control, i.e. a trend change. Trend changes are usually harder to identify and indicate a gradual loss of reliability. The warning limits are used to detect the gradual change in the average and should also include some decision criteria on how to handle this change. The decision criteria will depend on the criticality of the process.

Random errors are those which are caused by random and unpredictable variation in a process and may be seen as acceptable (with the normal variation of process) or unacceptable errors, i.e. those that fall outside the control limits.

2.4. Quality control of test systems

It is important that certain key go/no-go points are established during the preparation and use of the test system for an *in vitro* method. Key quality attributes (e.g., genetic/phenotypic stability, identity and absence of contamination), based on the suppliers' documentation and the facility's needs, and should be documented, with acceptance criteria, preferably in SOP(s). A QC plan to periodically confirm these attributes on a regular basis should be put in place. In practice it may not be always feasible to assess all "essential characteristics". The *in vitro* method should therefore include relevant and reliable positive and negative controls, including acceptance criteria, which will be used to establish an historic database of the test system essential functional characteristics. Lack of cell proper authentication, provenance, and characterisation could be grounds for a member country not accepting data that are not adequately documented.

Proprietary *in vitro* methods and the related *in vitro* systems may be relatively expensive; therefore their availability for QC testing may be limited by practical considerations, such as cost. In light of these considerations, the user may sometimes be dependent on the supplier to provide as complete as possible documentation regarding the test system, including cell or tissue characterisation and functional performance. The supplier should be expected to provide adequate documentation of quality control testing of each batch manufactured.

The suppliers' documentation should detail appropriate test system integrity checks of the Original Source (Table 2.1), ideally with evidence of test results provided by the supplier or a qualified service provider (Section 5.2.). These checks should also be performed on

the cells arriving in the laboratory as soon as samples can be obtained (Early Stocks). Ideally Cell Banks (both master and working) should be established ((Coecke et al., 2005[3]) but testing may be focused on the master stock with more routine checks applied to working cell banks e.g., mycoplasma and viability (Table 2.1). In addition, the user should carry out quality control checks in the test facility on a regular basis (Routine Culture testing) appropriate to the test system so that the *in vitro* method performs as expected after transport and handling of the test system.

Table 2.1. Applicability of integrity checks on cell cultures

Attributes	Original Source	Early Stocks	Cell Banks	Routine Cultures
Morphology	✓	✓	✓	✓
Viability	✓	✓	✓	✓ [a]
Identity	✓	✓	✓	
Doubling time [b]	✓	✓	✓	✓
Mycoplasma	✓	✓	✓	✓
Viruses	✓ (donor only)		✓ (master bank only)	
Bacteria and Fungi			✓	✓ [c]
Function/phenotype		✓	✓	✓ [d]
Genetic stability			✓	✓ [e]
Absence of reprogramming vectors (iPSC [f] lines)		✓	✓	

Notes:

[a] Viability testing at passage will also be helpful to ensure consistent seeding of fresh cultures and assays for more reliable maintenance of stock cultures and reproducibility of cell-based *in vitro* methods. For this, the assays described under Section 6.10.1 can be applied.

[b] For diploid cultures subcultured at a 1:2 ratio, passage number is roughly equal to the number of population doublings (or population doubling level) since the culture was started.

[c] To avoid development of low grade contamination, sterility testing may be desirable for long term cultures. These may also be sustained as separate replicate sets of flasks to provide backup cultures in case of contamination.

[d] Assessed by the correct performance of reference/control items.

[e] A risk/benefit analysis should determine if genetic stability analysis is required e.g. pluripotent stem cells.

[f] Induced Pluripotent Stem Cells (also known as iPS cells or iPSCs) are a type of pluripotent stem cell that can be generated directly from adult cells. The iPSC technology is based on the introduction of specific genes encoding transcription factors that can convert adult cells into pluripotent stem cells.

Where primary cell cultures and tissues are used, variation in properties between individual donors must be considered, and each new batch should be qualified or controlled for key functionality (Meza-Zepeda et al., 2008[4]); Special care should be taken to note any unusual observations in case of contamination or viral cytopathic effects or transformation, and all primary cell cultures should ideally be cryopreserved and screened for mycoplasma. Human and animal tissues and primary cells used for testing will also need to be appropriately documented. As part of QC for tissues, their differentiated state should also be documented, which may require a range of assays including for instance morphology, histochemistry, cell markers, specific tissue function and cell-cell/matrix interactions (Stacey and Hartung, 2006[5]). For primary cells prepared from tissues stored as banks of cryopreserved vials of cells, similar QC approaches can be used as adopted for banks of continuous cell lines (Section 5.5.).

Moreover, records recommended by Good Cell Culture Practice (GCCP) or other relevant guidance documents (e.g., ISO standards, GLP) should be kept. Guidance on cell and tissue culture work is available for either general (Coecke et al., 2005[3]) or specific applications (Andrews et al., 2015[6]; Geraghty et al., 2014[7]; ISCBI, 2009[8]; Pamies, 2016[9])).

2.5. Quality control of consumables and reagents

All consumables and reagents should be evaluated to be fit for the intended purpose(s). Consumables such as flasks, cryovials, culture dishes, culture slides, tubes, cell scrapers, etc. in general will not require any in-house QC, however it is good practice to maintain any relevant documentation provided by the supplier, such as proof of sterility, date of arrival, expiry dates and batch numbers, as the suitability and acceptability of materials may be questioned by the GLP Monitoring Authorities.

Test facilities can perform quality control checks of consumables, but the process how to do this is not always evident. Some test facilities have established procedures whereby a percentage of consumables from each batch/lot number are evaluated prior to use in *in vitro* work (e.g., for sterility testing). While this approach will not prevent contamination, it can provide data which can be useful for future evaluation of contamination. Ideally, sterile consumables with appropriate certificates should be used where possible. Alternatively, some consumables can be treated with ultraviolet (UV) light, gamma irradiation and/or autoclaved. Viral infection via such biological material as Foetal Calf Serum (FCS) (Section 4.3.1) can be avoided e.g., by gamma ray radiation of FCS (House, House and Yedloutschnig, 1990[10]; Nuttall, Luther and Stott, 1977[11]). These preventive measures may be useful in limiting contamination. Other consumables, such as centrifugal filter units and filtered pipette tips, cannot be pre-treated. In the case where no commercial sterile centrifugal filter units and/or filtered pipette tips are available, establishing a method for detecting contamination from these items is very important.

Certain materials which are critical to the performance of a method may be subject to significant variation, such as growth promoting reagents, hormones and conventional serum products (functional tests including acceptance criteria need to be defined). These critical reagents should be reliably available and sourced from a reputable supplier (where possible alternative sources should be identified), and should either be accompanied with the supplier's Certificate of Analysis (CoA), or appropriate quality controls should be applied in-house (Good Cell Culture Practice (GCCP) and Good Cell Culture Practice for stem cells and stem cell derived models). These controls may include growth or functional characterisation and should be performed by qualified personnel according to documented procedures or formal SOPs.

For some critical reagents it may be necessary to test for batch to batch variability so as to reduce the introduction of unknown variables, which may interfere with assay or overall *in vitro* method performance. For this purpose a batch is tested first and when approved, a large quantity of the batch can be acquired to reduce variability during the performance of a certain number of assays. Successive batches may be tested in-house and the new batches compared against historical data (e.g., growth rates).

For established reagents, the *in vitro* method uses the acceptance criteria of negative controls to identify eventual issues related to a new batch of reagents. Similar reagents obtained from different suppliers may each have specific and not necessarily the same

acceptance criteria. Acceptance criteria should be established for reagents depending on the degree of risk they represent to the final results. This risk can be assessed by:

1) Considering the potential impact of the perceived risk to prioritise certain reagents.
2) Formally evaluating the Quality Management System (QMS) of the supplier and establishing suitable Agreements (e.g., Service Level Agreements (SLA)) with the provider ensuring quality, availability and shipment of the reagent. Acceptance of individual batches of reagents can be addressed by review of key elements of the certificate of analysis, compliance with specific conditions of the agreement provided by the manufacturer/supplier or a combination of these and supplementary evaluation which may include pre-use testing to assure that individual batches are fit for purpose.
3) Assessing consistency of batch/lot qualification tests on critical reagents.

2.6. Staff training and development

Training is an integral part of all quality assurance systems, and must be formally planned and documented. For example GLP requires the maintenance of records of qualifications, training, experience and job descriptions of personnel (OECD, 1998[12]). Training should be formal, approved (certified), documented to a standard format and typically described in a SOP (WHO, 2009[13]). Training should be proactive, enabling staff to acquire the skills and knowledge that, with experience, makes them competent in the cell and tissue culture aspects of their work or enables them to elicit an appropriate reactive response where necessary. New objectives and new activities or procedures (e.g., SOPs) will always involve some training, and therefore requires new certification of the involved personnel. GLP attaches considerable importance to the qualifications of staff, and to both internal and external training given to personnel, in order to maintain levels of excellence and ensure the procedures are performed consistently by all personnel.

A list of core training for *in vitro* cell culture laboratory staff is detailed in the GCCP (Coecke et al., 2005[3]). Special aspects of training are also referred to in other sections of this document where relevant.

Documented training plans are useful to define procedures in which staff should be trained before they are considered competent. Regular review by line managers of staff performance is a useful tool for considering ongoing training needs. These may include regulatory requirements (e.g., GLP training), specific *in vitro* methods and their associated proficiency chemicals, use and storage of documentation, as well as general training in best practice such as indicated in GCCP (Coecke et al., 2005[3]). When new staff is recruited to work in the laboratory, it is important to guide the staff and review and document any training requirements before assignment to carry out any tasks. It may be helpful to demonstrate competence by documenting individual elements of training followed, including competence to perform the procedure(s) independently.

It is good practice to record all training in individual training files, including training records and periodic competency reviews. Supplementary training and education should also be documented to demonstrate maintenance of ongoing professional development to provide assurance that current best practices are maintained.

2.7. Types of documentation

The importance of documentation cannot be over stressed as it is the only way to demonstrate the work performed, i.e. if it is not documented it did not happen. It should enable reconstruction of a study/experiment and is also essential for the interpretation of the results.

Documentation in a quality system typically involves documents and records at several levels. The main document is a high-level, accurate description of the types of work performed by the organisation or group, key policies and standards adopted for delivering the work and the structure of the quality system. In some systems, this may be called a "quality manual". Another level may include overviews of procedures referring to the various specific testing methods involved at the next level. Finally, supporting the SOPs, there will be formal record sheets for test and control data and templates for reporting results. An overview of descriptive and prescriptive documents is provided in the World Health Organisation (WHO) handbook on quality practices in Biomedical Research (WHO, 2013[14]).

The WHO divides documentation into two broad classes:

- Prescriptive documents that give instructions as to what is to happen during the course of a study, such as SOPs and Study Plans.
- Descriptive records that describe what actually happened during the course of the study, such as records of raw and derived data, study reports.

Many quality systems require document management to assure that all documents are developed and approved in a formal process, that versions are accurately dated, authored and approved with specific version numbers to avoid inadvertent use of obsolete documents.

Each institution should implement rules regarding the recording and retention of data. Record keeping, whether by hand or making entries to electronic systems, should meet certain fundamental elements of data integrity (Section 10.1).

2.8. Quality considerations regarding electronic data integrity

The integrity of electronic data, and how to assess it, should be described in the quality system. Some of the common issues that repeatedly come up in US Food and Drug Administration (FDA) warning letters are:

- **Common passwords.** Sharing of passwords, or use of common passwords, does not allow the true identification of the operator, i.e. it is not attributable.
- **User privileges.** The software application is not adequately configured so as to define or segregate user levels and users may have access to inappropriate privileges such as modification of methods and integration or even deletion of data.
- **Computer system control.** Access to the operating system is not adequately implemented and users may modify system configurations (e.g., system clock) or allow unauthorised access to modify or delete electronic files; the file, therefore, may not be original, accurate, or complete.

- **Processing methods.** Integration parameters are not controlled and there is no procedure to define integration leading to concerns over re-integration of chromatograms at a later time.
- **Audit trails.** In this case, the laboratory has turned off the audit-trail functionality within the system. It is, therefore, not clear who has modified a file or why.

See Section 10.1 for a more in-depth discussion on data integrity.

References

Andrews, P. et al. (2015), "Points to consider in the development of seed stocks of pluripotent stem cells for clinical applications: International Stem Cell Banking Initiative (ISCBI)", *Regenerative Medicine*, Vol. 10/2s, pp. 1-44, http://dx.doi.org/10.2217/rme.14.93. [6]

Coecke, S. et al. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33/3, pp. 261-287. [3]

Geraghty, R. et al. (2014), "Guidelines for the use of cell lines in biomedical research", *British Journal of Cancer*, Vol. 111/6, pp. 1021-1046, http://dx.doi.org/10.1038/bjc.2014.166. [7]

House, C., J. House and R. Yedloutschnig (1990), "Inactivation of viral agents in bovine serum by gamma irradiation", *Canadian Journal of Microbiology*, Vol. 36/10, pp. 737-740, http://dx.doi.org/10.1139/m90-126. [10]

ISCBI (2009), "Consensus Guidance for Banking and Supply of Human Embryonic Stem Cell Lines for Research Purposes", *Stem Cell Reviews and Reports*, Vol. 5/4, pp. 301-314, http://dx.doi.org/10.1007/s12015-009-9085-x. [8]

Meza-Zepeda, L. et al. (2008), "High-resolution analysis of genetic stability of human adipose tissue stem cells cultured to senescence", *Journal of Cellular and Molecular Medicine*, Vol. 12/2, pp. 553-563, http://dx.doi.org/10.1111/j.1582-4934.2007.00146.x. [4]

Nuttall, P., P. Luther and E. Stott (1977), "Viral contamination of bovine foetal serum and cell cultures", *Nature*, Vol. 266/5605, pp. 835-837, http://dx.doi.org/10.1038/266835a0. [11]

OECD (2016), *Application of Good Laboratory Practice Principles to Computerised Systems*, OECD Series on Principles on Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing Paris. [1]

OECD (2016), *OECD Position Paper Regarding the Relationship between the OECD Principles of GLP and ISO/IEC 17025*, OECD Series on Principles of GLP and Compliance Monitoring, No. 18, OECD Publishing, Paris. [2]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [12]

Pamies, D. (2016), "Good Cell Culture Practice for stem cells and stem-cell-derived models", *ALTEX*, http://dx.doi.org/10.14573/altex.1607121. [9]

Stacey, G. and T. Hartung (2006), "Availability, Standardization and Safety of Human Cells and Tissues for Drug Screening and Testing", in *Drug Testing in vitro*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, http://dx.doi.org/10.1002/9783527609611.ch9. [5]

WHO (2013), *WHO Handbook: Quality practices in basic biomedical research*, World Health Organization. [14]

WHO (2009), *Handbook: Good Laboratory Practice (Second Edition) - Quality practices for regulated non-clinical research and development*, World Health Organization. [13]

Chapter 3. Facilities

Key message: *In vitro cell and tissue culture facilities should be fit for purpose and a detailed understanding of the work flow for the in vitro method related processes is essential. The separation of specific laboratory functions and elements that can adversely impact in vitro method work need to be understood.*

Key content: *Elaborates safety, safety risk assessment and management including descriptions of Risk Groups and Biosafety Level requirements, proper facility design to ensure integrity of the cell and tissue cultures, the in vitro method itself and the resulting data.*

Guidance for improved practice: *This chapter describes guidance on level of separation to avoid cross-contamination and quarantine measures for new test systems. A flow diagram indicating movement of staff, materials and reagents, test systems and test and reference items, and waste collections shows what processes need to be separated.*

Recommendations *for classification of infective microorganisms, laminar flow biological safety cabinets and biosafety levels, are given.*

Facilities must be fit and suitable for the purpose of the work; that is, size, construction, and location should be appropriate, and the building should allow for the separation of activities.

Higher containment levels may be required depending on the biosafety risk level (Section 3.2.2) of the biological agents handled. If *in vitro* work is to be performed with test systems belonging to Risk Group III or IV (Section 3.2.1), which can cause severe human disease and may be a serious hazard to employees or spread to the community, then separated facilities, appropriate Biosafety Levels (BSLs) such as air filtration and negative pressure differences, will need to be maintained (Section 3.2.2). Risk Groups III and IV are more complex in complying with specific facility requirements and personnel skills. Therefore if possible, *in vitro* methods for regulatory use in human safety assessment should be developed to require mainly BSL 2 or less.

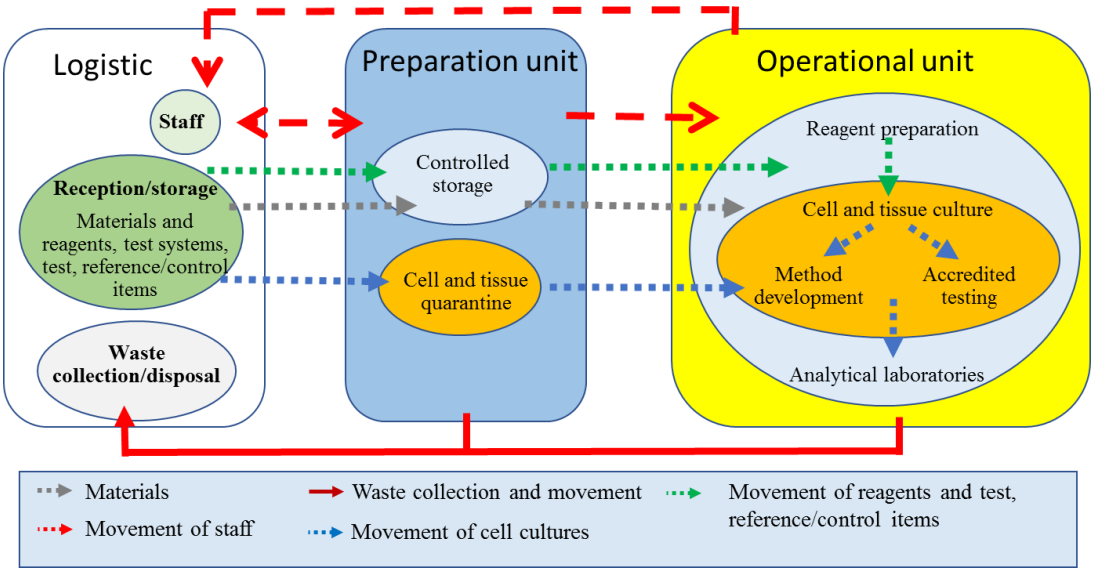
3.1. Facility design

When designing a new facility or modifying an existing facility, a safety risk assessment should be performed (Section 3.2), as safety should be included in the design phase. This is more critical for BSL 3 or 4 facilities (Section 3.2.2). It is important to understand the workflow for the intended processes and those aspects which could impact adversely on others, so that the facility design will facilitate smooth and safe laboratory procedures, storage and waste disposal.

Facilities should be designed or adapted to minimise the risk of errors (e.g. mix-ups) and to avoid (cross-)contamination which may adversely affect the quality of the work performed. Services (e.g., gas, electricity, liquid nitrogen) should ideally be accessible for routine maintenance to minimise interference in laboratory work. All the necessary permits should be in place before any activities are initiated. Finally, there should be dedicated areas for data storage and archiving. An Uninterruptible Power Supply (UPS) should also be available for all critical equipment, including reagent/sample storage[1] (especially critical for low-boiling point reagents) to ensure preservation in case of loss of power.

The types of laboratory functions, along with the flow of work and materials, outlined in Figure 3.1 are among those to be taken into consideration for separation (physical or process/training), that someone establishing or running a facility should be aware of. It may not be possible or acceptable to separate all functions. Other functions, specific to the type of work performed, may also need to be taken into consideration. It is wise to avoid physical contact between materials transfers and waste removal so that there is very low risk of contamination from waste affecting reagents, cultures and test materials.

Figure 3.1. Flow of staff and materials to show separation of processes



Note: It is recommended that each area have their own dedicated storage facilities so as to avoid mixing up test items and/or reagents.

As contaminated working surfaces can lead to microbial contamination or cross-contamination between test systems and pose a risk to the *in vitro* method quality, working surfaces should be easy to clean, resistant to acids, alkalis, solvents and disinfectants. There should be appropriate documented procedures for disinfection of work surfaces, safety cabinets and equipment.

Physical separation of pre and post Polymerase Chain Reaction (PCR) assay stages should be maintained to minimise contamination and cross-contamination (Section 3.3). Between these two areas the work flow should be unidirectional. Equipment, consumables and laboratory coats should each have a dedicated area. It is recommended that facilities performing PCR methods should be organised in four discrete areas[2]. Requirements may vary with assay format e.g., real time PCR does not require post-PCR analysis[3].

1. **Reagent preparation clean area** - air pressure should be positive.
2. **Nucleic acid extraction area** - air pressure should be positive. If chemicals are stored in this area appropriate facilities and storage requirements should be in place.
3. **Amplification area** - PCR machines are housed in the Amplification room. It may contain an area/cabinet with air pressure slightly positive for the nested PCR.
4. **Product analysis area** - air pressure should be negative.

Good Laboratory Practice (GLP) test facilities require archive(s) that should provide for the secure storage and retrieval of study plans, raw data, final reports, samples of test items and specimens (OECD, 2004[1]). Archive design and archive conditions should protect contents from untimely deterioration.

3.2. Safety, risk assessment and management

Countries or regions may have specific classification of microorganisms and/or other hazards, which should be consulted when performing the safety risk assessment.

A risk management approach should be used when introducing new processes or when modifying the design of the facility, so as to eliminate potential hazards prior to their introduction. Risk management is a process or method used to identify, evaluate and determine the appropriate way to deal with exposure to hazards and risk factors that have the potential to cause harm, and is an ongoing process that requires continuous review to ensure that the implemented control measures work as planned.

When planning a safety risk assessment all hazards should be considered, including physical, chemical, photic and biological hazards and should comply with all national and/or international legislation. The risk assessment should not be limited to just the laboratory, but should also consider the entire site and any possible risks to the environment, including waste disposal for any hazardous materials and again should comply with national laws.

Safety risk assessments should be performed by the individual(s) most familiar with the specific characteristics of the test systems being considered for use, the equipment, materials and reagents, the procedures to be employed, and the containment equipment and facilities available. Exposure to these hazards might be complex and may require specialist knowledge both in identifying and evaluating their associated risks and designing appropriate actions to avoid or minimise them.

The transport of dangerous items should also be addressed in the risk assessment, specifically what precautions to take in case of spillage. International transport, either by rail, air or road should comply with international norms, e.g., International Air Transport Association (IATA) and/or the Dangerous Goods Regulations (DGR).

Training of staff (Section 2.6) in preventative procedures such as the correct use of Biological Safety Cabinets (BSCs), aseptic techniques, use of personal protection equipment (PPE), waste disposal, etc., will not only ensure a safer working environment but will also benefit the quality of the work performed.

3.2.1. Risk Groups

In many countries biological agents are categorized in Risk Groups Figure 3.1 based on their relative risk. Most countries have national or local laws and regulations governing safety in the workplace. Many of these national regulations classify microorganisms based on the biological risks they present to human health and/or to the environment. While no agreed international classification scheme exists, the WHO formulated a set of minimum standards for laboratory safety detailing four risk groups, the last version having been published in 2003. Variation of these four risk groups have been implemented into national laws worldwide.

Table 3.1. Classification of infective microorganisms by risk group

Risk Group I	low individual and community risk	A microorganism that is unlikely to cause human disease or animal disease of veterinary importance.
Risk Group II	moderate individual risk, limited community risk	A pathogen that can cause human or animal disease but is unlikely to be a serious hazard to laboratory workers, the community, livestock, or the environment. Laboratory exposures may cause serious infection, but effective treatment and preventive measures are available and the risk of spread is limited.
Risk Group III	high individual risk, low community risk	A pathogen that usually produces serious human disease but does not ordinarily spread from one infected individual to another.
Risk Group IV	high individual and community risk	A pathogen that usually produces serious human or animal disease and may be readily transmitted from one individual to another, directly or indirectly.

Source: (WHO, 2004[2])

3.2.2. Biosafety Levels

The assignment of a biosafety level should take into consideration a multitude of factors, such as the microorganism or pathogenic agent used, the facilities available, the equipment practices and procedures and should not be just the automatic assignment according to the particular risk group (WHO, 2004[2]).

BSL prescribes procedures and levels of containment for the test systems and materials. Test facilities may be assigned to one of four BSL based on a safety risk assessment (Table 3.2).

Table 3.2. Summary of Biosafety Level requirements

	1	2	3	4
Isolation[a] of laboratory	No	No	Yes	Yes
Room sealable for decontamination	No	No	Yes	Yes
Ventilation				
• inward airflow	No	Desirable	Yes	Yes
• controlled ventilating system	No	Desirable	Yes	Yes
• HEPA-filtered air exhaust	No	No	Yes/No[b]	Yes
Double-door entry	No	No	Yes	Yes
Airlock	No	No	No	Yes
Airlock with shower	No	No	No	Yes
Anteroom	No	No	Yes	-
Anteroom with shower	No	No	Yes/No[c]	Yes
Effluent treatment	No	No	Yes/No[c]	Yes
Autoclave				
• on site	No	Desirable	Yes	Yes
• in laboratory room	No	No	Desirable	Yes
• double-ended	No	No	Desirable	Yes
Biological safety cabinets	No	Desirable	Yes	Yes
Personnel safety monitoring capability[d]	No	No	Desirable	Yes

Source: (WHO, 2004[2])

Cell lines and primary tissues may carry a variety of different microorganisms or pathogens, which can potentially cause human disease, pose hazard to employees and

distort the *in vitro* method results. These should be handled at biosafety (hazard) level 2, unless known to be pathogen free including any likely serious human pathogens based on the origin and species of the material. This level of containment is also appropriate for monoclonal antibody-containing supernatants and cell homogenates. Access to level 2 facilities should be restricted to authorised personnel only, and specific safety risk assessment and training activities should be followed according to the national legislation on level 2 containment (Coecke, 2005[3]; Geraghty et al., 2014[4]).

3.2.3. Biological safety cabinets

BSCs are designed to use HEPA filters to capture particles 0.3 micron or bigger. Many are also designed to recirculate a high percentage of the filtered air within the cabinet (e.g., Class II A1 and A2 recirculate about 70% of the filtered air in the BSC) and to exhaust the remaining filtered air into the room or, if fitted, to an exhaust system. When handling toxic chemicals in a BSC it is critical to know the percentage of air exhausted into the room, due to the possibility of personnel exposure, and also the percentage recirculated in the BSC as it could potentially affect the test system and therefore the results generated with the *in vitro* method.

Chemical fume hoods, on the other hand, are designed to capture, contain, and exhaust hazardous fumes generated inside the hood, and should be used to handle and prepare hazardous chemicals whenever possible. When handling highly toxic materials the use of a glove box may be preferred (consult the facility's chemical risk assessment and/or the suppliers safety data sheets for the correct handling requirements). If a hazardous chemical is to be used in a BSC, the BSC should be equipping with an active carbon filter on the hood exhaust. The quantity of hazardous chemical used must be limited and where possible the pure hazardous chemical should not be handled inside the BSC.

Most commercially available BSCs are certified (e.g., the public health and safety organization NSF International/The American National Standards Institute ANSI 49 - 2012, Biosafety Cabinetry: Design, Construction, Performance, and Field Certification, BS/EN 5726 Microbiological safety cabinets) for the stated classification. This certification should also be confirmed once the cabinet has been installed in the facility so as to guarantee its proper functioning and regular testing performed to assure correct ongoing function.

A commonly accepted classification based on their containment capabilities and performance attributes has been adopted by most manufactures (Table 3.3). At a minimum, all cell and tissue work should be performed in a Class II biological safety cabinet as even screened tissues or cell cultures may carry infectious agents not covered by virological screening. Class I BSC, where the airflow is directed inward into the cabinet, provides protection for personnel and the environment but not for materials or work inside the cabinet, as it does not prevent contact with airborne contaminants that may be present in laboratory air. Class II BSC, often referred to as vertical laminar flow cabinets due to a unidirectional HEPA-filtered air stream that descends downward, provide protection for the personnel, the environmental and for the work performed inside the cabinet (Table 3.3). For guidance in use of Class II cabinets refer to the Good Cell Culture Practice (GCCP) principles (Coecke, 2005[3]).

Table 3.3. Classification of laminar flow biological safety cabinet

Classification	Biosafety Level	Protection Provided	Application
Class I	1, 2, 3	Personnel and Environmental Protection Only	low to moderate risk biological agents
Class II	1, 2, 3	Product, Personnel and Environmental Protection	low to moderate risk biological agents
Class III	4	Total Containment Cabinets	high risk biological agents

If microscopes or other equipment are to be installed in a BSC, the cabinet should be checked for flow disruption so as to maintain the correct functionality of the BSC.

Splashes and aerosols carry contamination and infection risks, which not only endanger the operator but may also compromise the integrity of the *in vitro* method (i.e., cross contamination of cell lines or introduction of adventitious agents) results. Therefore, all procedures should aim at minimising aerosol production. Any procedures likely to produce aerosols should be contained (e.g., using a BSC) or the material should be rendered harmless.

3.2.4. Waste disposal

Prior to introducing new or modifying existing procedures it is necessary, and often required by law, to carry out a safety risk assessment which will include the assessment of any potential risks related to the waste generated. For most commercially acquired chemicals and reagents the suppliers' documentation will enable rapid assessment of potential associated risk. For test systems acquired from commercial cell providers the provided documentation may also be used to facilitate the risk assessment; however, for test systems obtained from another laboratory the documentation provided, if any, will rarely be sufficient, placing an extra burden on the facility. It is important that decontamination procedures are also put in place and are tested for their efficacy against those microorganisms likely to be present. Laboratory generated waste should be disposed of on a regular basic and not allowed to build up in the facility. The flow of waste removal within the facility should be such as to minimise potential secondary contamination.

3.3. Strategies to avoid cross-contamination

It is the responsibility of all laboratory staff to ensure that the correct workflow is followed and appropriate training should be given to the personnel regarding the necessary precautions to minimise contamination and cross-contamination, e.g., training on the use of aerosol-free/aerosol filter pipet tips when working with PCR assays.

Measures should be taken to ensure adequate separation of different biological agents and studies taking place in the same physical environment (OECD, 2004[1]). The integrity of each test system and study should be protected by spatial or temporal separation from other studies and test systems to avoid potential cross-contamination and mix-up. The flow of materials, staff and waste can be an important factor in controlling these issues and Figure 3.1 gives an illustration of how this may be applied.

Tissues and cells from different studies can be kept in the same incubator provided that they have the same incubation temperature requirements, are labelled appropriately, are spatially separated and none of the test items or solvents used are volatile enough to cause

contamination. Tissues and cells from different species or *in vitro* methods where yeast and bacteria are used require a higher level of separation. The most important issue here is to separate the areas used for cell culture/tissue and microbiological culture and that adequate care (e.g., use separate protective clothing) is taken not to carry over contamination from one area to the other, which would ideally be described in a SOP. Other degrees of separation may be achieved using the specific requirements described elsewhere for quarantine of untested material.

Temporal separation of test systems is possible in biological safety cabinets by handling only one test system at a time. Before introducing a new test system the cabinet working surfaces and related equipment should be cleaned and decontaminated, for example by cleaning with 0.5% solution of hypochlorite (approx. 5000 ppm free chlorine) followed by 70% isopropyl alcohol and then wiping with sterile wipes. The cabinet may then be exposed to UV light, if appropriate.

Rooms and areas used for preparation and mixing of test and reference items with vehicles should allow for aseptic working conditions in order to minimise the risk of microbial contamination of the test system.

When performing molecular biology techniques and especially PCR-based assays, which are high sensitivity methods, extreme care should be taken in facility design (Section 3.1) and operation. False-positive results can originate, for example, from sample-to-sample contamination from carry-over of nucleic acid from previous amplification of the same or similar target. Cloned DNA or virus-infected cell cultures may represent other source of contamination[4].

A major source of PCR contamination is aerosolised PCR products (Scherczinger et al., 1999[5]). Once these aerosols are created, being small, they travel and easily spread all over benches and equipment, where they can find their way into a PCR reaction and become amplified. Laboratories exclusively performing real-time PCR and properly discarding all amplified products without opening the reaction tubes or using sealed plates are less liable to contamination and could therefore be dispensed from the follow-up measures.

A no template control and a reverse transcription negative control should always be included in the PCR reaction test runs to exclude contaminations in reagents, in the work environment etc. When performing real time PCR, the use of dUTP in place of dTTP in the dNTP mix is recommended, in this way, all amplicons generated will have dUTP incorporated in them. In the future, if that amplicon becomes the source of contamination, using the enzyme Uracil-DNA-glycosylase prior to PCR specifically targets dUTP-containing DNA, resulting in excision of uracil, and prevents PCR contamination by a previous amplicon. The excision of uracil prevents the amplicon amplification by creating abasic sites in the amplicons. The abasic sites do not serve as good DNA templates for Taq polymerase. Therefore, the contaminated amplicons are prevented from being amplified further (Nolan et al., 2013[6]; Taylor et al., 2010[7]).

It is recommended to colour code racks, pipettes and laboratory coats in the different areas so to be able to easily monitor their movement between the different areas. Powder-free gloves should be used throughout the process in all the different areas as the powder on powdered gloves might affect the assay outcome/performance. It is particularly important to always use powder-free gloves in the pre-PCR area, as the pre-PCR area is prone to contamination by RNases.

The reagent preparation clean room should be free from any biological material such as DNA/RNA, cloned material, etc. Primers and reagents aliquoting is recommended to minimise contamination consequences. To ensure clean areas are kept free of amplicon at all times, there should be no movement from the dirty area to the clean area. If under extreme circumstances a consumable or reagent needs to be moved back it must be thoroughly decontaminated with bleach and ethanol. Returned racks should be soaked in a 0.05% solution of hypochlorite overnight before soaking in distilled water and placing in the clean area. To ensure minimal movement between areas during the running of molecular assays, it is optimal to have dedicated storage (freezer, fridge and room temperature) for each area. Room air pressure should be positive.

In the **nucleic acid extraction room/area** samples are processed, reverse transcriptase step of RT-PCR are performed and DNA or cDNA, and positive controls are added to the PCR reaction mix (prepared in the reagent preparation clean room).

Post-PCR manipulations such as agarose gel electrophoresis are performed in the P**roduct analysis room/area.** It is thus a contaminated area and therefore no reagents, equipment, pipettes, coats, etc. used in this room should be used in any other PCR areas. Bench areas should be wiped daily with hypochlorite solution following use and contaminated areas should be additionally decontaminated with ultra-violet radiation if available. Hypochlorite solutions containing more than 500 ppm available chlorine are corrosive to some metals and alloys and should not be applied to stainless steel (types 304/347, 316 and 400 series) as it may lead to corrosion with repeated use. It is recommended for personnel working with post-PCR assay stages not to work with pre-PCR parts later the same day. Monitoring of viable and non-viable particles of critical equipment surfaces and air flow within these areas/rooms may also be beneficial in controlling contamination.

3.4. Air handling, water supply, environmental control, heating and cooling

Air handling systems should be operated to ensure that the correct environment is maintained for the type of work conducted in the laboratory. These systems should be subject to regular maintenance and serviced by qualified personnel and records of maintenance, including modifications, should be retained to demonstrate appropriate upkeep and function. Where the *in vitro* work involves the use of human pathogens, the laboratory should operate with specific trained personnel, using biosafety level 3 or 4 and the room should be kept at negative pressure to guard against infection spread. When High-Efficiency Particulate Air (HEPA) filters are used in differential pressure isolation rooms, the filters and their fittings and seals need to be thoroughly examined and tested at regular intervals (e.g., annually). Decontamination should be carried out before servicing is carried out. Air handling systems should also be designed to account for exhaust air from the Class II biological safety cabinets to be vented outside.

Cell culture work requires cell/tissue culture grade water, which is usually deionised using reverse osmosis membrane separation, followed by passage through a series of carbon and micropore filters eliminating organic materials and pyrogens. Tissue culture grade water should be controlled for pH, conductance and total organic carbon. Note that pyrogens can be deleterious to cell cultures at concentrations below the level of detection for organic carbon. Where small quantities of purified water are required for cell culture, sterile Water For Injection[5] (WFI) or other medically approved pure-water preparations may be used (Stacey and Davis, 2007[8]).

Heating, cooling and humidity should be adequate for the comfort of laboratory occupants and for operation of laboratory equipment, and should not adversely affect test system survival/behaviour and test item stability. For example, in some cases (e.g., preparation of microscopic slides) specific humidity might be required. Desiccation of cell culture media should be avoided and most modern incubators will have humidification systems installed as standard.

Many tissue culture media components are sensitive to white light (especially sun light), with blue wavelengths being of particular concern. Filters can be used in the room, on the windows and laminar flow cabinet light to reduce this exposure where necessary.

Mid to long term storage of media is usually best at temperatures below ambient laboratory temperatures. Accordingly, an optimal solution may be to store all cell culture media at 2°C to 8°C (refrigerator) or frozen (freezer) or as recommended by the manufacturer. There may be exceptions to this general rule but the manufacturers' instructions should always be consulted.

3.5. Quarantine for new test systems

New cells and tissues should be quarantined in the laboratory or in storage until determined to be free of contaminating microorganisms (Figure 3.1). However, exceptions may be made for specific cases: e.g., human blood samples cultured for chromosomal aberration test cannot be stored and must be used on the second day after receipt. It is important for those cases where quarantine is not possible to have supplier documentation, e.g., CoA indicating freedom of contamination. There may also be some cases where the CoA or proof of freedom of contamination is not provided directly with the test system, e.g., some 3D tissue. In these cases the contamination aspects should be assessed in parallel with the work and all work performed in a controlled environment. The test facility should not release any data acquired with this test system until freedom of contamination has been proven. Regular tests to identify contamination of microorganisms during the subsequent cell and tissue culture life cycle, including cell banking, are recommended (Section 5.7).

Early checks of cell authentication (Section 5.6) are also recommended to avoid wasted time and resources on unauthentic cell lines. If separate laboratories/hoods/incubators are not available, steps should be taken to minimise the risk of spreading contamination (Geraghty et al., 2014[4]). Alternatively, other steps can be taken to minimise contamination risks, such as handling the quarantine cells last on each day, rigorous post-manipulation disinfection of the work areas and placing cultures for incubation in a filter-sealed container into the general incubator (Geraghty et al., 2014[4]) (Any area used for the handling of quarantined materials should be routinely cleaned after each use, using a suitable disinfectant. Cells procured from a cell bank may be accompanied with a certificate of analysis which may list the contamination checks performed and provide details of testing methods. At a minimum, a mycoplasma test (Section 5.9, table 9) should be performed upon receipt and cell cultures carefully observed for evidence of contamination.

Notes

1. Refers to prepared samples (e.g., cells treated with test, reference or control items)
2. Separate rooms or containment areas (such as PCR workstation, laminar flow cabinet).
3. See: https://www.gov.uk/government/publications/smi-q-4-good-laboratory-practice-when-performing-molecular-amplification-assays
4. See: https://www.gov.uk/government/publications/smi-q-4-good-laboratory-practice-when-performing-molecular-amplification-assays
5. Also known as Water for Irrigation (WFI)

References

Coecke, S. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33, no. 3, pp. 261-287. [3]

Geraghty, R. et al. (2014), "Guidelines for the use of cell lines in biomedical research", *British Journal of Cancer*, Vol. 111/6, pp. 1021-1046, http://dx.doi.org/10.1038/bjc.2014.166. [4]

Nolan, T. et al. (2013), *Good practice guide for the application of quantitative PCR (qPCR)*, https://www.gene-quantification.de/national-measurement-system-qpcr-guide.pdf. [6]

OECD (2004), *The Application of the Principles of GLP to in vitro Studies*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 14, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084971-en. [1]

Scherczinger, C. et al. (1999), "A Systematic Analysis of PCR Contamination", *Journal of Forensic Sciences*, Vol. 44/5, p. 12038J, http://dx.doi.org/10.1520/jfs12038j. [5]

Stacey, G. and J. Davis (eds.) (2007), *Medicines from Animal Cell Culture*, John Wiley & Sons, Ltd, Chichester, UK, http://dx.doi.org/10.1002/9780470723791. [8]

Taylor, S. et al. (2010), "A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines", *Methods*, Vol. 50/4, pp. S1-S5, http://dx.doi.org/10.1016/j.ymeth.2010.01.005. [7]

WHO (2004), *Laboratory biosafety manual Third edition*, World Health Organization, http://www.who.int/csr/resources/publications/biosafety/Biosafety7.pdf?ua=1. [2]

Chapter 4. Apparatus, material and reagents

Key message: *Apparatus, including validated computerised systems, should be regularly maintained, calibrated and validated (if required). Material and reagents should be purchased from well-established sources to ensure the integrity and reliability of the in vitro method results.*

Key content: *Quality requirements for equipment, material and reagents (e.g., use of serum, alternatives to the use of animal sourced serum, antibiotics, special media, certificate of analysis, stability and traceability) are detailed.*

Guidance for improved practice: *By detailing the diversity in availability of in vitro related materials and reagents the reader can identify for his/her work their advantages and limitations.*

Recommendations *are given to reduce experimental variability and increase within- and between-laboratory reproducibility by understanding the material and reagents you are working with and to take care that calibrated apparatus performance checks are carried out and operation limits are set adequately.*

4.1. Apparatus

Apparatus, including computerised systems, used for the generation, storage and retrieval of data, and for controlling environmental factors relevant to a study should be suitably located and of appropriate design and adequate capacity. Apparatus should be periodically inspected, cleaned, maintained, and calibrated according to Standard Operating Procedures (SOPs) and records of these activities should be maintained (OECD, 1998[1]). In general, all apparatus used should be operated by trained staff.

The routine requirements for apparatus used in a Good Laboratory Practice (GLP) environment apply equally to apparatus used for *in vitro* studies (OECD, 2004[2]). However to ensure the integrity and reliability of some results, certain equipment such as microbalances, plate readers, centrifuges, micropipettes, laminar air flow biological safety cabinets, fridges and freezers, water baths, and incubators should be regularly maintained, monitored and calibrated (if applicable). Calibration standards should be traceable to international standards if possible. For each type of equipment, critical parameters (e.g., supply of gases for mass spectrometry, liquid nitrogen levels in storage containers, low temperature storage (fridge/freezers), temperature and CO_2 levels in incubators) should be identified as requiring continuous monitoring or the setting of limit values together with installation of alarms.

Centrifuges which are routinely used in cell and tissue culture work (subculturing, cryopreservation, etc.) may produce aerosols and therefore it is important to consider models that have sealed buckets. Ideally, one should consider working with models where the condition of the load can be observed without opening the lid. Besides the containment issues for centrifuges it is necessary to specify centrifugation speeds as G-force (g) rather than Revolutions Per Minute (RPM) (unless the rotor radius is stated), incubation conditions, time and volumes of centrifugation with tolerances when relevant, and any other information that enables the accurate reproducibility of procedures. In addition, procedures should be established on cleaning, including cleaning frequency of buckets, caps, adapters, rotor, and bulkhead so as to reduce the possibility of contamination of cultures. Procedures should also be established regarding potential exposure and how to respond in case of an emergency (e.g., broken tubes).

Working with cell and tissue culture requires a strictly controlled environment for cell growth. This is achieved using specialised incubators which provide the correct growth conditions (temperature, humidity, CO_2 levels), which should be controlled (and logged) on a regular basis. Incubators that use a nebuliser to deliver humidity are preferred to older models which use a water pan/basin for the same purpose. This combination of high humidity and temperature increases the risk of bacterial or fungal contaminations and therefore care should be taken when using a water pan/basin equipped incubator. If using an incubator with a water pan/basin, sterile distilled (or equivalent) water should be used, and antifungal or bactericidal agents can be added to the water pan/basins to reduce the risk of bacterial and fungal growth. However, any possible impact due to the use of these agents on the *in vitro* method to be carried out should be checked and documented. Good practice is to avoid contact of any bactericidal/fungicidal agent with the cells and/or reagents used in tissue or cell culture. Another option to reduce the risk of microbial contamination is to use copper-coated incubators which are now available. Incubators with self-sterilising cycles may also be used, although this does not replace regular cleaning and maintenance.

Similarly, water baths used to thaw and/or to warm up stored solutions like medium and frozen stocks, or to defrost vials of cryopreserved cells and tissues, carry a high risk of introducing contamination. Sterile or deionised water should be used and the water should be regularly changed. It is good practice to carefully wipe down media bottles and/or cryopreserved vials with paper towels wetted with 70% (isopropyl) ethanol or other sterilising solutions before their transfer to a Biological Safety Cabinet (BSC). The use of bactericidal and fungicidal agents in water baths can aid in the control of contamination, but their impact on the test system should be checked and documented, and avoided where possible. Bead baths may also be used so as to reduce cross-contamination that may be more likely in water baths, especially when using tubes that may not be water-tight. Bead baths sometimes take longer to get up to the set temperature and accidental spills or contamination requires thorough washing and decontamination of the beads.

A BSC (Section 3.2.3) should be considered as a critical piece of equipment for cell and tissue culture work, since, when it is used correctly (Section 3.2.3), it ensures a clean working environment providing protection for both the operator and for cells/tissues and other materials and reagents. BSCs require regular service and maintenance such as integrity testing of High-Efficiency Particulate Air (HEPA) filters, testing of airflow velocity profile and testing of non-viable particle counting to make sure the cabinet is fully functional. Laboratory personnel must be fully trained in how to work within the BSC so to maintain aseptic culture technique.

For equipment such as refrigerators and freezers, temperatures should be checked regularly and preferably logged, e.g., using data loggers to record the temperature at set intervals. In addition to the regular recording of temperatures, an alarm system to alert staff when acceptable operating limits are exceeded is desirable, and a backup system should be in place, such that materials may be transferred from one fridge/freezer to another, in case of malfunction or for cleaning.

Acceptable operating limits should be set, monitored and recorded for all measuring equipment. Equipment should be fit for purpose with respect to sensitivity and selectivity. Equipment used to perform measurements should be calibrated[1] or verified[2], usually described in the facility SOP(s), at specified intervals or prior to use. As an example, pipettes or micropipettors may need to be checked more frequently than centrifuges. If during the checking errors are encountered, the pipette may need to be adjusted and recalibrated to ensure it meets the stated acceptable operating limits. A maintenance schedule should be implemented detailing the frequency of maintenance (e.g., yearly) of all equipment.

When pipetting volatile/viscous liquids or suspensions, it is strongly recommended to use positive displacement pipettes. Certain chemicals may exhibit non-specific adsorption to the plastic tips of pipettes and the use of low-binding materials (including glass) or acoustic droplet ejection (Ekins, Olechno and Williams, 2013[3]; Grant et al., 2009[4]) and can be utilised to alleviate these issues.

It may be necessary to have separate procedures for regular checks and complete calibration depending on the frequency of use and the criticality of the instrument. The frequency of checking may be extended if historical data shows low failure rates. When equipment such as a pipette is out of specifications during a calibration, it is important to determine how to interpret data that have been generated since the most recent successful calibration and determine the impact of the potential deviation on the outcome of the study. Therefore, it is crucial to record every piece of equipment, uniquely identified, that has been used during the performance evaluation of an *in vitro* method. In general,

facility practices should ensure that equipment is within specifications before the start of a study and throughout the experimental phase to avoid rejection of the *in vitro* study data.

Complex instrumentation, i.e. computerised systems including robotic systems, should be formally validated prior to use in a GLP study and procedures should be established to ensure that these systems are suitable for their intended purpose and are operated and maintained in accordance with the Principles of Good Laboratory Practice (OECD, 2016[5]; OECD, 1998[1]). The level of validation will depend on the systems complexity and its intended use and usually includes documentation of User Requirement Specifications (URS), a validation plan and report, user acceptance testing and reporting. More complex systems require in addition formally documented qualification of the system via Design Qualification (DQ), Installation Qualification (IQ), Operational Qualification (OQ), and Performance Qualification (PQ), where the IQ and OQ may be performed by the supplier/manufacturer. Whatever approach is taken, it should be justified by a documented risk assessment (FDA , 2001[6]; OECD, 2016[5]). The Official Medicines Control Laboratories (OMCLs) guidelines for qualification of equipment may be helpful in designing the validation for a given computerised system (OMCL, 2011[7]).

To enable broader use of a new method, successful transfer to a range of equipment (if applicable) and different laboratories should be demonstrated. This increases the robustness of the method. To increase transferability, preference should be given to the use of generally widely available equipment. In addition, the impact of the use of a certain type or brand of equipment on the outcome of the individual assays or the overall *in vitro* method needs to be determined. An *in vitro* method should specify the requirements the equipment should meet to be used for the specific method.

4.2. Materials and reagents

Reagents are often selected on the basis of historical use or from references in relevant documents associated with regulatory accepted *in vitro* methods (e.g., validation reports, *in vitro* method SOPs). It is good practice to have procedures for maintaining and controlling laboratory stocks of reagents such as maintaining a minimum stock level for critical reagents.

The *in vitro* method should use reagents from well-established sources (to avoid as much as possible labour intensive control checks), preferably certified suppliers. OECD recommends that suppliers implement the International Standard ISO 9001, and particularly Part 1 - Specification for Design/Development, Production, Installation and Servicing (OECD, 2000[8]). Nowadays most suppliers have adopted manufacturing practices which comply with formal national or international standards, such as ISO 9001. Identification and qualification of alternate suppliers for critical reagents and materials is also recommended.

Preparation of reagents should be tracked (e.g., by use of logbook or electronic record(s)) to retain information such as the supplier, catalogue number, batch/lot numbers (if appropriate), dates of preparation and expiry, and the names of the operator involved in the preparation. For both reagents and reagents mixtures, the container should be inert and not affect the stability of the substance or mixture. Attention will also need to be given to the suitability of reagents and to the safety and ethical provenance of cells (Coecke et al., 2005[9]; Pamies et al., 2016[10]).

Labelling of reagents should be defined in a procedure (SOP) and should include identity, concentration (if appropriate), expiry date and specific storage instructions. The expiry date may be extended on the basis of documented evaluation or analytical analysis (OECD, 1998[1]).

Storage should be done according to the manufacturer's specifications as detailed in the Certificate of Analysis (CoA) or product information sheet. Some solutions, e.g., solutions which require storage below 0°C, may be aliquoted in order to minimise the number of times a bottle is opened and thus minimise the risk and spread of contamination and avoid repeated freeze/thaw cycles. When reagents need to be thawed and possibly frozen again, it is recommended to determine the number of freeze/thaw cycles that the reagents can withstand (EMA, 2011[11]; FDA , 2001[6]; Viswanathan et al., 2007[12]); (FDA , 2001[6]). Stability of aliquots should be verified in the laboratory performing the *in vitro* method and should not be based solely on literature data.

Even when reagents are sourced from a reputable supplier, it remains important to assure the stability of the reagents during shipment conditions, in addition to the storage. For example, reagents shipped frozen should arrive frozen and this should be documented on the receiving document. The presence of a data logger is the best practice in these cases.

Quality checks, if required, should be performed according to pre-defined procedures described in SOPs. Normally, stability of the analyte in the studied matrix is evaluated using at least triplicate samples of the low and high concentrations, which are analysed immediately after preparation and after the applied storage conditions that are to be evaluated. The thawed samples are analysed against a calibration curve, obtained from freshly prepared calibration standards, and the obtained concentrations are compared to the nominal concentrations. The deviation should be within previously established acceptance criteria (usually ±20%) (EMA, 2011[11]). It is absolutely necessary that the number of cycles in the freeze/thaw stability evaluation should equal or exceed that of the freeze/thaw cycles of study samples.

4.3. Use of media in cell culture

Depending on the circumstances, the basal culture medium can be animal serum-supplemented (as in traditional cell culture methods) or serum-free, but supplemented with additives necessary for obtaining satisfactory cell proliferation and production, or for maintaining a desired differentiation status. Many slightly different formulations exist under the same general medium names, such as Minimum Essential Medium (MEM), and even subtle changes in the medium formulation can substantially alter the characteristics of certain cells and tissues. In many cases, these variations are deliberate for specific applications. Therefore, the medium to be used should be precisely specified, and it is essential to check that new supplies of medium meet the required specifications (Coecke et al., 2005[9]). If a medium other than that recommended/indicated by the cell provider is used then the justification should be documented and the effect on baseline cell properties should be determined and provided with the final data.

4.3.1. The use of animal sourced serum in cell culture

Serum is a complex mixture, introducing undefined components into the medium. Many of these substances have not yet been identified, and in many cases the effects on cultured cells are as yet unclear. Animal serum can be derived from adult, new born or foetal sources, but typically less than 24 months old animal sources should be used[3] (Festen,

2007[13]). Bovine sera are most commonly used and Foetal Calf Serum (FCS)[4] has become the standard supplement for cell culture media in the last few decades.

In vitro method developers must determine serum specifications that meet their particular needs and match the natural behaviour of the cells as much as possible, including defining the maximum acceptable levels of serum components, such as immunoglobulins (which may have inhibitory effects), endotoxins (indicative of bacterial contamination, but are also powerful cell mitogens), and haemoglobin (indicative of haemolysis during clotting). Furthermore, if the quality of the serum is deemed critical to the performance of the method, more rigid testing requirements will apply and should be specified in the respective Test Guideline (TG).

Ideally, sera should be obtained from vendors that can provide traceability certification from industry bodies such as the International Serum Industry Association (ISIA). Vendor's documentation, usually in the form of a CoA, generally include country of origin and traceability information, filtration steps used in serum processing, sterility testing, screening for mycoplasma and virus, endotoxin, lot number, storage conditions etc. (Sadeghi et al., 2017[14]). Test facilities rely on the documentation the supplier provides, including the compatibility of different lots/batches of serum.

Batches of serum can differ dramatically in their ability to support the growth of cell lines due to variation in the concentration of growth factors and hormones, therefore, new batches should be tested on the appropriate cell line(s) for cell attachment, spreading, cloning efficiency, growth rates and activity in functional assays (Geraghty et al., 2014[15]). Testing of serum batches will ensure in-house reproducibility. Some facilities, based on experience with specific test systems, do not always perform full additional batch testing however this should be judged on a case by case basis.

Serum can interfere with phenotypic cell stability, and may influence experimental outcomes. Serum can suppress for instance embryonic stem cell differentiation and tissue formation. The use of FCS can possibly lead to unexpected or undesired outcomes, e.g., FCS can inhibit transforming growth factor (TGF)-β1-induced chondrogenesis in fibroblast-like type-B synoviocytes (Bilgen et al., 2007[16]). FCS compared to autologous (human) serum has been found to induce a more differentiated and less stable transcriptional profile in human bone marrow mesenchymal stem cells, particularly at late passages, as shown by analysis of genome-wide microarray analysis (Shahdadfar et al., 2005[17]).

Cell lines which have been derived or cultured long-term in serum-containing media may become dependent on the multitude of growth factors present and may experience a phenotypic drift upon abrupt serum withdrawal. This may manifest as growth arrest or activation/inactivation of various signalling pathways. These effects can be overcome by adaption to serum-free culture conditions (Section 4.3.3) using specific protocols (Beltran et al., 2014[18]; Leong et al., 2017[19]; Sinacore, Drapeau and Adamson, 2000[20]) for a gradual weaning of cells (van der Valk et al., 2010[21]).

4.3.2. The use of animal sourced serum in cell culture for endocrine activity

To study the effects of chemical substances that may have endocrine activity (e.g., steroid hormones), endogenous hormones, growth factors and cytokines are removed by charcoal stripping of serum. If FCS is required in Endocrine Active Substances (EAS) *in vitro* methods, it is necessary to use Dextran-Coated-Charcoal-treated Foetal Calf Serum[5] (DCC-FCS) when performing these tests. DCC-FCS as a basic component of cell culture

medium has become the standard supplement and has been listed in several OECD TGs, e.g., TG 455 and TG 458.

A 2005 study found that DCC-FCS affected the commitment of osteoprogenitor KS483 cells, strongly promoting adipogenesis compared to normal FCS containing medium, which drives KS483 cells to differentiate into only osteoblasts (Dang and Lowik, 2005[22]). This suggests possible unpredictable effects of DCC- FCS on progenitor cell differentiation.

4.3.3. Alternatives to the use of animal sourced serum

The use of serum has been discouraged in recent years due to the undefined nature of the medium, batch variability that may contribute to experimental variability and lack of reproducible data, and potential limitation in consistency and availability of supply. Moreover, *in vitro* methods, including components, are often developed for legislative or ethical reasons to replace animal methods. In 2008 the ECVAM Scientific Advisory Committee (ESAC) stated that "*for methods forwarded to ECVAM for validation/prevalidation where [the use of non-animal alternatives to serum] is not fulfilled a justification for future use must be provided, including measures taken to seek non-animal alternatives to [FCS]*" [6]. The drawbacks of using FCS and the recommendation to replace it with available chemically defined serum free media is already discussed in the GCCP guidance document issued by EURL ECVAM (Coecke et al., 2005[9]). Furthermore, it is recommended to develop new *in vitro* methods with a serum-free, chemically-defined medium, to avoid potential sources of uncertainty that may be introduced by using animal serum (Jochems et al., 2002[23]; Pamies et al., 2016[10]).

Serum-free media (Table 4.1) are thought to circumvent many of the drawbacks of using FCS including the batch to batch variability issues associated with serum and offer better reproducibility and the potential for selective culture and differentiation of specific cell types (Geraghty et al., 2014[15]). Nevertheless, serum-free compositions may still need to be validated and monitored similarly to serum containing media as they are often not completely chemically defined. A source of a range of commercially available serum-free media for cell-culture, as well as medium compositions obtained from scientific literature, is provided by the 3Rs-Centre ULS in collaboration with Animal Free Research UK (FCS-free database[7]).

Table 4.1. Serum-free media

Media	Description
Serum-free	Does not require supplementation with serum, but may contain discrete proteins or bulk protein fractions (e.g., animal tissue or plant extracts) and are thus regarded as chemically undefined (see: chemically defined media).
Protein-free	Does not contain high molecular weight proteins or protein fractions, but may contain peptide fractions (protein hydrolysates), and are thus not chemically defined. Protein-free media facilitate the down-stream processing of recombinant proteins and the isolation of cellular products (e.g., monoclonal antibodies), respectively.
Animal-product-free	Does not contain components of animal or human origin. These media are not necessarily chemically defined (e.g., when they contain bacterial or yeast hydrolysates, or plant extracts).
Chemically defined	Does not contain proteins, hydrolysates or any other components of unknown composition. Highly purified hormones or growth factors added can be of either animal or plant origin, or are supplemented as recombinant products (see: animal-product-free media).

Source: (van der Valk et al., 2010[21])

The use of human serum was originally restricted to specialised applications (Coecke et al., 2005[9]). However due to better quality controls, including documentation to demonstrate origin and viral safety, human serum has become more widely used (Blázquez-Prunera et al., 2017[24]; Dessels, Potgieter and Pepper, 2016[25]; Even, Sandusky and Barnard, 2006[26]; Gstraunthaler, Lindl and van der Valk, 2013[27]); (Dessels, Potgieter and Pepper, 2016[25]; Even, Sandusky and Barnard, 2006[26]; Gstraunthaler, Lindl and van der Valk, 2013[27]; Jochems et al., 2002[23]; Kanafi, Pal and Gupta, 2013[28]); and has been shown feasible by adapting the KeratinoSens™ skin sensitisation test to xeno-free cell culture (Belot et al., 2017[29]). The same critical points, e.g. batch-to-batch variability, as for any serum-derived products hold true.

Human platelet lysates (hPLs) have been proposed as an alternative growth supplement to FBS. hPLs are the result of freeze-thawing platelet concentrates and contain several growth factors (Bieback et al., 2009[30]). Platelet concentrates, typically products manufactured for transfusion purposes, can be used as a cell culture supplement after the shelf life of the donation program has expired. As these programs are managed by certified blood donation centres, hPLs are therefore obtained from safe and clinically tested sources. hPLs have now been successfully used in several applications such as growth and maintenance of renal epithelial cell lines and human mesenchymal stromal cells, and storage of human tissues for patient related treatment (van der Valk, 2018[31]). hPLs cannot be considered a defined supplement, though.

Other serum free media can include poorly defined supplements such as pituitary extracts, chick embryo extracts, bovine milk fractions or bovine colostrum. Furthermore, some so-called 'defined' media contain complex serum replacement mixtures including chemically undefined agents. Notably B27 and its alternative NS21 used in the culture of neural cells contain bovine serum albumin and transferrin which can exhibit batch to batch variation in biological activity (Chen et al., 2008[32]). Therefore, it may be useful to carry out pre-use testing on new batches of reagents which could demonstrate variability that cannot be foreseen from the manufacturers' information. Another example of an essential component prone to batch to batch variability is the so-called 'basement membrane extract', purified from Engelbreth-Holm-Swarm (EHS) mouse sarcoma cells and marketed under various trade names.

Chemically-defined media are cell-type specific, in contrast to FCS (van der Valk, 2018[31]), and have to be selected and optimised for the selected cell type (Price, 2017[33]) These media are commercially available for many cell types[8], but the formulations of these are generally not released because of proprietary reasons. Therefore, it is important to check before use what cell lines a chemically-defined medium is available or optimised for. It is also possible to develop a chemically-defined medium and adapt cells as defined by van der Valk (van der Valk et al., 2010[21]). Applying factorial design approaches have been shown to minimise the screening time, allow prediction for best medium formulation and can be used as a high-throughput medium optimisation platform (Zhao et al., 2017[34]).

Serum-free medium formulations for culturing of stem cells, such as human Embryonic Stem Cells (hESCs) and human induced Pluripotent Stem Cells (hiPSCs), show promise for applications in toxicology, regulatory testing and biomedical research (Colatsky et al., 2016[35]). A recent study (van Velthoven et al., 2017[36]) indicates that stem cells *in vivo* may have a very different gene expression profile *in vitro* which should be taken into consideration when conducting *in vitro* studies examining stem cell function. Both hESCs and hiPSCs are often maintained on inactivated mouse embryonic fibroblasts or under

feeder-free conditions (using extracellular matrices) in chemically defined, serum-free media, in order to avoid the presence of undefined or unknown serum components (which may compromise the differentiation towards desired cell lineages) and the risk of contaminations from pathogens (e.g., mycoplasma, viruses, and prions) (Pistollato et al., 2012[37]; Yamasaki et al., 2014[38]).

4.3.4. The use of phenol red in cell culture

Phenol red is used in the cell culture as a convenient way to rapidly check on the health of cell or tissue cultures. In the initial cell culturing stage, a small amount of phenol red is often added to the cell culture medium. Under normal conditions most living cells or tissues prosper at a near-neutral pH, i.e. a pH close to 7, and the culture medium has a pink-red colour as an indicator colour. Under abnormal conditions, cellular waste products or contaminants will cause a change in pH, leading to a change in indicator colour.

Phenol red can interfere with some spectrophotometric and fluorescent assays, and it is also weakly estrogenic. To avoid the possibility of interference with specific assays, it is therefore recommended that phenol red-free medium be used.

4.4. The use of antibiotics in cell culture

Routine cell and tissue culture according to GCCP (Coecke et al., 2005[9]; Geraghty et al., 2014[15]; Stacey and Davis, 2007[39]) should not require the use of antibiotics as it can never be relied on as a substitute for effective aseptic techniques. However, its use is still widespread e.g., OECD TG 432 (OECD, 2004[2]) due to established routine procedures in many laboratories. Antibiotics are agents that may arrest or disrupt fundamental aspects of cell biology, and, while they are effective against prokaryotic cells (i.e. bacteria), they are also capable of causing toxic effects in animal cells. Not surprisingly, antifungal agents, being directed at higher order, eukaryotic microorganisms, are likely to be more toxic to animal cell cultures. In addition, antibiotics often make it more difficult to detect microbial contamination. Given these obvious contra-indications, the use of antibiotics in cell and tissue culture should be focused in two areas: a) protection of materials at high risk of contamination such as tissues, organs and primary cultures in cases where sterility cannot be guaranteed; and b) the positive selection of recombinant cell clones based on the expression of antibiotic resistance genes (Coecke et al., 2005[9]). If antibiotics are needed, a justification for the use of antibiotics in the procedure is recommended.

4.5. Additional media components

Some media components are heat labile (e.g., L-glutamine), sensitive to light (e.g., retinoic acid) or have a limited half-life in diluted state or at high ionic strength, such as in prepared media (e.g., epidermal and fibroblast growth factors). These issues are best addressed by preparing a small volume of media necessary to cover the period of stability of the most sensitive component and discarding bottles after a set time period. Appropriate size aliquots of those labile components may be frozen by an appropriate method for long-term storage. In this respect, stock solutions with low concentrations of protein aqueous growth factors may require the addition of albumin or other excipients to prevent adsorption to plastic and to increase stability in the frozen state. Stabilised forms of glutamine and retinoic acid are also available to avoid these issues.

In case culture media or other reagents have to be sterilised via heat or filtration, the impact of the procedure (e.g., comparison of doubling time to historical data) should be assessed and recorded. For example, heat sterilisation may result in degradation (or denaturation) of one or more of the components and filtration can remove individual and/or essential components (e.g., Fe^{2+} or Fe^{3+} iron products that enhance growth of mammalian eukaryotic cells in serum-free cultures).

4.6. Dedicated media for particular cell lines

Different cell types or tissues need to be cultured in media containing various components at different concentrations to allow optimal growth. Although certain cell lines may be grown in media with the same composition, sharing media between cell lines increases the risk of cross-contamination. Therefore, each cell line should be cultured with separate dedicated media, which must not be shared with other cell lines. It is important to note that different media types are not only used for different cell cultures, but also for the same culture (e.g., when differentiating HepaRGs to hepatocyte-like cultures or primary lung epithelium cells in air-liquid interface culture different media are used in the differentiation procedures). Sharing media between laboratory personnel also increases the risk of contamination and cross-contamination and should be avoided.

Notes

1. Calibration, as used here, is a measurement against a known standard and may involve adjustment of the apparatus, which may or may not be described in the equipment manual.
2. Verified, as used here, is a confirmation that the device fulfils specified requirements where no adjustment is possible.
3. http://www.thermofisher.com/it/en/home/life-science/cell-culture/mammalian-cell-culture/fbs/other-sera/bovine-serum.html
4. Also known and available as Foetal Bovine Serum (FBS)
5. Also known as Dextran-Coated-Charcoal-treated Foetal Bovine Serum (DCC-FBS)
6. https://eurl-ecvam.jrc.ec.europa.eu/about-ecvam/archive-publications/publication/ESAC28_statement_FCS_20080508.pdf
7. https://fcs-free.org/
8. https://fcs-free.org/

References

Belot, N. et al. (2017), "Adaptation of the KeratinoSens™ skin sensitization test to animal-product-free cell culture", *ALTEX*, pp. 560-564, http://dx.doi.org/10.14573/altex.1701311. [29]

Beltran, P. et al. (2014), "Adaptation to serum-free culture of HEK 293T and Huh7.0 cells", *BMC Proceedings*, Vol. 8/Suppl 4, p. P259, http://dx.doi.org/10.1186/1753-6561-8-s4-p259. [18]

Bieback, K. et al. (2009), "Human Alternatives to Fetal Bovine Serum for the Expansion of Mesenchymal Stromal Cells from Bone Marrow", *STEM CELLS*, Vol. 27/9, pp. 2331-2341, http://dx.doi.org/10.1002/stem.139. [30]

Bilgen, B. et al. (2007), "FBS suppresses TGF-β1-induced chondrogenesis in synoviocyte pellet cultures while dexamethasone and dynamic stimuli are beneficial", *Journal of Tissue Engineering and Regenerative Medicine*, Vol. 1/6, pp. 436-442, http://dx.doi.org/10.1002/term.56. [16]

Blázquez-Prunera, A. et al. (2017), "Human mesenchymal stem cells maintain their phenotype, multipotentiality, and genetic stability when cultured using a defined xeno-free human plasma fraction", *Stem Cell Research & Therapy*, Vol. 8/1, http://dx.doi.org/10.1186/s13287-017-0552-z. [24]

Chen, Y. et al. (2008), "NS21: Re-defined and modified supplement B27 for neuronal cultures", *Journal of Neuroscience Methods*, Vol. 171/2, pp. 239-247, http://dx.doi.org/10.1016/j.jneumeth.2008.03.013. [32]

Coecke, S. et al. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33/3, pp. 261-287. [9]

Colatsky, T. et al. (2016), "The Comprehensive in Vitro Proarrhythmia Assay (CiPA) initiative — Update on progress", *Journal of Pharmacological and Toxicological Methods*, Vol. 81, pp. 15-20, http://dx.doi.org/10.1016/j.vascn.2016.06.002. [35]

Dang, Z. and C. Lowik (2005), "Removal of serum factors by charcoal treatment promotes adipogenesis via a MAPK-dependent pathway", *Molecular and Cellular Biochemistry*, Vol. 268/1-2, pp. 159-167, http://dx.doi.org/10.1007/s11010-005-3857-7. [22]

Dessels, C., M. Potgieter and M. Pepper (2016), "Making the Switch: Alternatives to Fetal Bovine Serum for Adipose-Derived Stromal Cell Expansion", *Frontiers in Cell and Developmental Biology*, Vol. 4, http://dx.doi.org/10.3389/fcell.2016.00115. [25]

Ekins, S., J. Olechno and A. Williams (2013), "Dispensing Processes Impact Apparent Biological Activity as Determined by Computational and Statistical Analyses", *PLoS ONE*, Vol. 8/5, p. e62325, http://dx.doi.org/10.1371/journal.pone.0062325. [3]

EMA (2011), *Guideline on bioanalytical method validation*. [11]

Even, M., C. Sandusky and N. Barnard (2006), "Serum-free hybridoma culture: ethical, scientific and safety considerations", *Trends in Biotechnology*, Vol. 24/3, pp. 105-108, http://dx.doi.org/10.1016/j.tibtech.2006.01.001. [26]

FDA (2001), *Guidance for Industry Bioanalytical Method Validation*. [6]

Festen, R. (2007), "Understanding Animal Sera: Considerations for Use in the Production of Biological Therapeutics", in *Medicines from Animal Cell Culture*, John Wiley & Sons, Ltd, Chichester, UK, http://dx.doi.org/10.1002/9780470723791.ch4. [13]

Geraghty, R. et al. (2014), "Guidelines for the use of cell lines in biomedical research", *British Journal of Cancer*, Vol. 111/6, pp. 1021-1046, http://dx.doi.org/10.1038/bjc.2014.166. [15]

Grant, R. et al. (2009), "Achieving Accurate Compound Concentration in Cell-Based Screening: Validation of Acoustic Droplet Ejection Technology", *Journal of Biomolecular Screening*, Vol. 14/5, pp. 452-459, http://dx.doi.org/10.1177/1087057109336588. [4]

Gstraunthaler, G., T. Lindl and J. van der Valk (2013), "A plea to reduce or replace fetal bovine serum in cell culture media", *Cytotechnology*, Vol. 65/5, pp. 791-793, http://dx.doi.org/10.1007/s10616-013-9633-8. [27]

Jochems, C. et al. (2002), "The use of fetal bovine serum: ethical or scientific problem?", *Alternatives to Laboratory Animals*, Vol. 30, pp. 219-227. [23]

Kanafi, M., R. Pal and P. Gupta (2013), "Phenotypic and functional comparison of optimum culture conditions for upscaling of dental pulp stem cells", *Cell Biology International*, Vol. 37/2, pp. 126-136, http://dx.doi.org/10.1002/cbin.10021. [28]

Leong, D. et al. (2017), "Evaluation and use of disaccharides as energy source in protein-free mammalian cell cultures", *Scientific Reports*, Vol. 7/1, http://dx.doi.org/10.1038/srep45216. [19]

OECD (2016), *Application of Good Laboratory Practice Principles to Computerised Systems*, OECD Series on Principles on Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing Paris. [5]

OECD (2004), *Test No. 432: In Vitro 3T3 NRU Phototoxicity Test*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264071162-en. [2]

OECD (2000), *Compliance of Laboratory Suppliers with GLP Principles*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 5, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078611-en. [8]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [1]

OMCL (2011), *OMCL Network of the Council of Europe Quality Assurance Document. Qualification of Equipment Core Document: PA/PH/OMCL (08) 73 2R*. [7]

Pamies, D. et al. (2016), "Good Cell Culture Practice for stem cells and stem-cell-derived models", *ALTEX*, http://dx.doi.org/10.14573/altex.1607121. [10]

Pistollato, F. et al. (2012), "Standardization of pluripotent stem cell cultures for toxicity testing", *Expert Opinion on Drug Metabolism & Toxicology*, Vol. 8/2, pp. 239-257, http://dx.doi.org/10.1517/17425255.2012.639763. [37]

Price, P. (2017), "Best practices for media selection for mammalian cells", *In Vitro Cellular & Developmental Biology - Animal*, Vol. 53/8, pp. 673-681, http://dx.doi.org/10.1007/s11626-017-0186-6. [33]

Sadeghi, M. et al. (2017), "Virome of US bovine calf serum", *Biologicals*, Vol. 46, pp. 64-67, http://dx.doi.org/10.1016/j.biologicals.2016.12.009. [14]

Shahdadfar, A. et al. (2005), "In Vitro Expansion of Human Mesenchymal Stem Cells: Choice of Serum Is a Determinant of Cell Proliferation, Differentiation, Gene Expression, and Transcriptome Stability", *Stem Cells*, Vol. 23/9, pp. 1357-1366, http://dx.doi.org/10.1634/stemcells.2005-0094. [17]

Sinacore, M., D. Drapeau and S. Adamson (2000), "Adaptation of Mammalian Cells to Growth in Serum-Free Media", *Molecular Biotechnology*, Vol. 15/3, pp. 249-258, http://dx.doi.org/10.1385/mb:15:3:249. [20]

Stacey, G. and J. Davis (eds.) (2007), *Medicines from Animal Cell Culture*, John Wiley & Sons, Ltd, Chichester, UK, http://dx.doi.org/10.1002/9780470723791. [39]

van der Valk, J. (2018), "Fetal bovine serum (FBS): Past – present – future", *ALTEX*, pp. 99-118, http://dx.doi.org/10.14573/altex.1705101. [31]

van der Valk, J. et al. (2010), "Optimization of chemically defined cell culture media – Replacing fetal bovine serum in mammalian in vitro methods", *Toxicology in Vitro*, Vol. 24/4, pp. 1053-1063, http://dx.doi.org/10.1016/j.tiv.2010.03.016. [21]

van Velthoven, C. et al. (2017), "Transcriptional Profiling of Quiescent Muscle Stem Cells In Vivo", *Cell Reports*, Vol. 21/7, pp. 1994-2004, http://dx.doi.org/10.1016/j.celrep.2017.10.037. [36]

Viswanathan, C. et al. (2007), "Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays", *Pharmaceutical Research*, Vol. 24/10, pp. 1962-1973, http://dx.doi.org/10.1007/s11095-007-9291-7. [12]

Yamasaki, S. et al. (2014), "Generation of Human Induced Pluripotent Stem (iPS) Cells in Serum- and Feeder-Free Defined Culture and TGF-β1 Regulation of Pluripotency", *PLoS ONE*, Vol. 9/1, p. e87151, http://dx.doi.org/10.1371/journal.pone.0087151. [38]

Zhao, A. et al. (2017), "Use of real-time cellular analysis and Plackett-Burman design to develop the serum-free media for PC-3 prostate cancer cells", *PLOS ONE*, Vol. 12/9, p. e0185470, http://dx.doi.org/10.1371/journal.pone.0185470. [34]

Chapter 5. Test systems

Key message: *With the advances in science and technology a variety of different cell and tissue culture-based test systems have been developed but only few have been used in regulatory-approved test guideline methods due to reliability issues caused by a variety of elements described in this chapter.*

Key content: *Elaborates on Good Cell Culture Practice logistics, cryostorage, handling, identification, containment, authentication and characterisation of the test system (e.g., cell lines, stem cells, primary cells, engineered tissues, etc.) already at the development stage.*

Guidance for improved practice: *Processes for checking test system identity and characteristics, comparison of ultra-low cryostorage methods and good subculture, cryopreservation and banking practices are given.*

Recommendations *are given for cell and tissue sourcing, contaminants screening, test system biomarkers and functional tests, since it may have influence on various aspects of the in vitro method.*

Data from *in vitro* cell and culture-based test systems are routinely used by industries and regulatory bodies in toxicity testing, safety assessment, and risk evaluation. The greatest use of *in vitro* test systems, however, is for elucidating mechanisms of toxicity and/or demonstrating the biological process involved, when exposing test systems to toxicants of various kinds.

In vitro methods utilise many types of test systems. The same biological source can be grown in different culture conditions, presented in different formats, and exposed to test item(s) through different means following different *in vitro* method procedures. For example, normal human keratinocyte cells can be cultured in a monolayer system for the neutral red uptake assay or cultured at the air liquid interface on a collagen matrix for the skin irritation test. These can be considered as two separate and distinct test systems and should be handled as such. Therefore, in this case it may be more appropriate to define the test system as the final preparation of those cells rather than normal human keratinocytes. As *in vitro* test systems become more sophisticated, the definition of the test system will need to cover the biological, chemical, or physical system in the finalised platform to be used for testing.

The need for more physiologically and human relevant *in vitro* test systems has led to a major effort to use microphysiological and microfluidic technologies in combination with advanced test systems including human stem cells (Watson, Hunziker and Wikswo, 2017[1]). With the advances in genetics and genetic screening approaches, routine *in vitro* methods already include the use of genetically altered cells, stem cells, stem-cell-derived models, organ-on-chip models (microphysiological systems; MPS) or other complex and sophisticated systems (Soldatow et al., 2013[2]). To date most of these novel methods are not yet ready for regulatory purposes, however rapid progress is being made with these new approaches.

The development process of such complex test systems requires characterisation in terms of viability, functionality, genotypic and phenotypic characteristics, which can be challenging. These extensive development efforts take place mainly in the *in vitro* method developer's laboratory. Moreover, reliability and performance of these novel *in vitro* methods will need to be determined before the method can be validated (Chapter 8).

5.1. Guidance on Good Cell Culture Practice

Good Cell Culture Practice (GCCP) identifies a set of core principles of best practice for working with simple but also with more complex cell and tissue culture systems (Good Cell Culture Practice (GCCP) and Good Cell Culture Practice for stem cells and stem-cell derived models). The principles of GCCP published in 2005 remain highly relevant to cell culture practice for *in vitro* methods today and may be applied to a broad set of applications, including research, manufacture of medicines, and laboratory based Good Laboratory Practice (GLP) testing. GCCP is a vital component of GIVIMP as it provides detailed and specific principles of best practice for the handling and management of cell and tissue culture systems.

As a result of a workshop organised in 2015 (Pamies, 2016[3]) scientists from European, Japanese and North American organisations identified new developments in cell and tissue cultures. The workshop report specifically addresses new technological developments in human pluripotent stem cell lines, stem-cell derived models and complex 3D cultures. Stem cells and their derivatives represent relevant *in vitro* toxicity models as they are characterised by unlimited self-renewal and the capacity to

differentiate into several human tissue-specific somatic cells such as liver cells, heart and brain cells.

5.2. Cell and tissue sourcing

A critical issue to consider when selecting a cell or tissue based test system is the source of the cells or tissues, as its history/handling may influence its characteristics and, consequently, the results of the *in vitro* methods conducted with this test system (Lorge et al., 2016[4]). Sourcing of cells and tissues from a certified provider, e.g., established cell banks with a high quality standard, commercial providers, or reputable culture collections (Table 1.1 Cell culture collections banks), who usually provide extensive documentation on the origins and characterisation of the test system is recommended[1]. If appropriate documentation is not provided, then each test facility will need to implement more rigorous processes for checking the identity and characterising the test system. Documentation of the absence of contamination by major classes of biological agents (e.g., mycoplasma, bacteria, fungi and viruses), genetic identity/consistency/traceability and stability of desired functionality should also be available. See Good Cell Culture Practice (GCCP) (Coecke et al., 2005[5]), GCCP principle 3[2] and Table 1.2 for examples of document requirements concerning the origins of cells and tissues.

Cell and tissue providers should be qualified by the test facility to assure appropriate documentation of cell and tissue origins and quality control key features (Section 2.4). An interesting example to mention is how the user community's joint efforts to define standardised cell sources in the field of genotoxicity made stocks of such mammalian cell lines available worldwide and issued recommendations for their handling and monitoring (Lorge et al., 2016[4]). In addition, the user should check that there is solid ethical provenance (e.g., the human Pluripotent Stem Cell Registry hPSCreg registry[3]) and safety assessment performed for the cells. Intellectual Property Rights (IPR) should also be checked to ensure that they do not impact on the use of the test system and future acquired data using the test system. For more detailed information on these issues see (Stacey et al., 2016[6]).

In the case of human tissues and primary cells, there is also a requirement to assure donor consent and to manage sensitive personal data. A broad range of issues in securing tissues for testing were addressed at the 32nd Workshop of the European Centre for Validation of Alternative Methods (Anderson et al., 1998[7]). Where tissues cannot be sourced via a qualified tissue bank, there should be an agreed testing method in place with clinical contacts regarding all aspects of harvesting, preparation, labelling, storage and transfer; for an example see (Stacey and Hartung, 2006[8]). It is also important to assess the risks of viral contamination of primary cells and tissues. More details on approaches for risk assessment of primary cells and tissues are described in (Stacey and Hartung, 2006[8]). Tissues should be obtained from tissue banks holding only materials from screened donors and this will significantly assist in managing viral safety issues. When working with human tissues and primary cells it is imperative to always follow national legislation.

Moreover, master and working cell banks, where applicable, should be established to guarantee a supply of constant quality and provide traceability to the original source (Section 5.5.1)

Where test systems used in *in vitro* studies are genetically modified, the Cartagena Protocol on Biosafety[4] provides a legal framework for international trade in Genetically

Modified Organism (GMOs) and provides Signatory State Parties with orientation and the framework for development of complementary national biosafety regulations (Bielecka and Mohammadi, 2014[9]). The Cartagena Protocol does not, however, address the risks and safe practices required when handling such organisms in the workplace. Therefore, specific measures for national implementation are necessary, e.g., Directive 2009/41/EC (EU, 2009[10]) in Europe.

5.3. Cell and tissue culture transportation

Many biological materials fall into the category of "dangerous goods" for shipping purposes and must comply with national regulations and/or international norms such as the International Air Transport Association (IATA) transport regulations[5] and/or the Dangerous Goods Regulations[6] (DGR). Diagnostic specimens of human or animal material including (but not limited to) blood and its components, tissue, tissue fluids or body parts are generally classified as Biological substance, Category B (UN3373[7]) for transport by air.

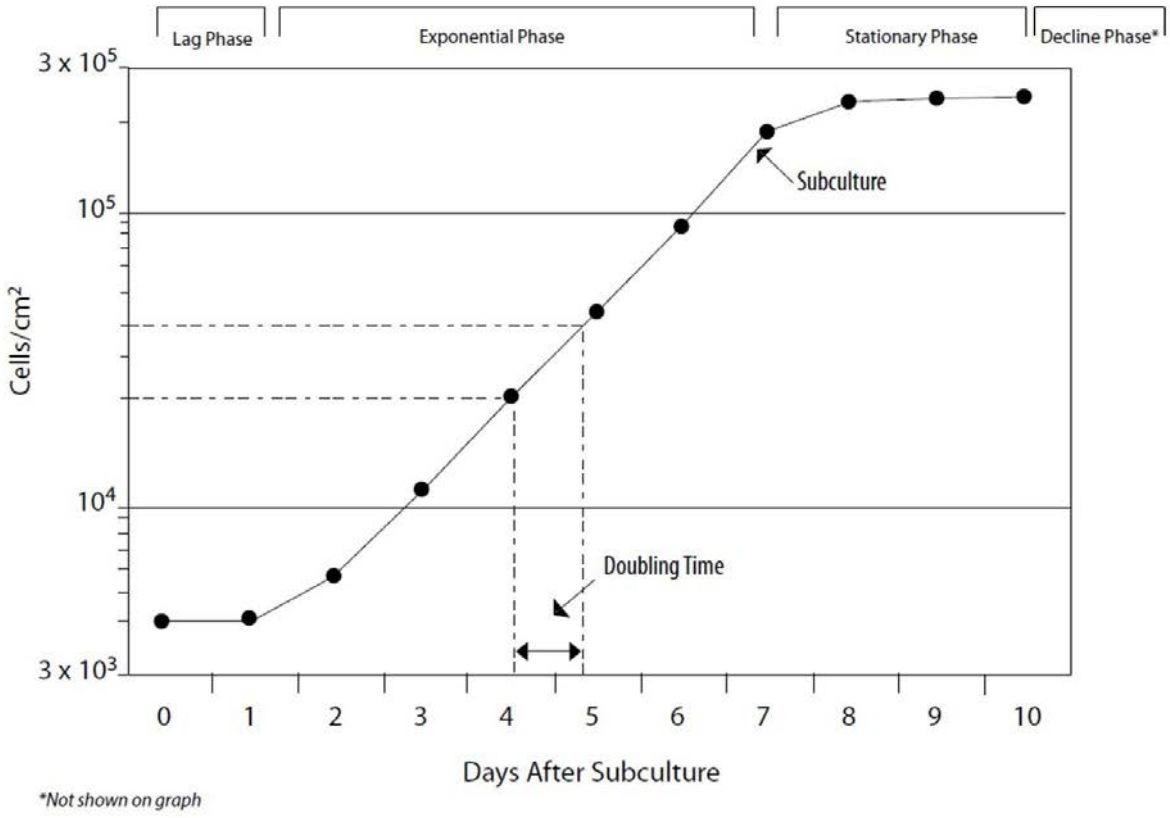
As cells and tissue in culture are often transported across the world, it is vital to keep these test systems as healthy as possible during the long transport times. Live cells and tissues may need to be shipped in a special temperature-controlled environment, such as that of a mini-cell culture incubator, where they are expected to reach their destination in good condition and are also less likely to become damaged during the transport process. Mini-cell culture incubators have limited space (2-3 plates or flasks) and require adequate sealing of plates to avoid leakage and may not always be an option or available. For short trips (e.g., arrive within one working day) it may be possible to ship cell lines in culture medium filled flasks.

Cell lines or cells are often shipped on dry ice. For shipment of some primary cells (e.g., primary liver cells) containers equilibrated with liquid nitrogen are used. Ideally, temperature should also be monitored (e.g., by using data-logger) during transportation, especially for long distance transport. When cells arrive at their destination, the conditions of the cell and tissue cultures should be examined and documented. Care should be taken when planning the shipment that the package does not sit over the weekend which may possibly compromise the test system integrity. Extra precautions must be taken for international shipping, as there is the possibility of samples being held up at customs. The fastest shipping times should be selected when long distances are involved.

5.4. Handling and maintenance of the test system

During routine handling and maintenance, growth and survival characteristics of the cell system (such as cell viability, doubling time, etc.) and subculturing details (e.g., date of subculture, subculture intervals, morphology, seeding density, passage number, etc.) should be recorded and documented, since they are required for the complete traceability of results. The documentation provided by the test system supplier (Table 1.2) should be taken into account together with historical data, when available, and used to establish acceptance criteria.

Figure 5.1. Growth curve for cells grown in culture



Source: (ATTC, 2014[11])

Figure 5.1 shows a typical growth curve for cells grown in culture. Whether cells grow and divide in a monolayer or in suspension, they usually follow the same characteristic growth pattern in which four different phases can be recognised: (a) a lag phase where the cells adapt to the new conditions; (b) a log or exponential phase of fast growth; (c) a plateau or stationary phase after the cells have completely covered the growing surface (are confluent) or saturated the suspension culture and (d) a decline phase where the cells begin to die. In order to ensure viability, genetic stability, and phenotypic stability, cell lines should ideally be maintained in the log or exponential phase, i.e. they need to be subcultured before a monolayer becomes 100% confluent or before a suspension reaches its maximum recommended cell density. The biochemistry of confluent/saturated cells may also be different from that of exponentially growing cells, and therefore, for most purposes cells are harvested or passaged before they become confluent or saturated. Some cell cultures can remain as confluent or saturated cultures for long periods, whereas others tend to deteriorate when they reach confluence. Some cell lines, particularly those derived from normal tissues such as human diploid fibroblasts, may be contact inhibited at confluence (Riss and Moravec, 2004[12]).

Two terms are commonly used to track the age of a cell culture: (i) passage number - indicates the number of times the cell line has been sub-cultured and (ii) the population doubling level[8] (PDL) - indicates the number of times a two-fold increase (doubling) in the total number of cells in culture has occurred since its initiation *in vitro*. PDL is not

determined for continuous cell lines as they are passaged at higher split ratios (ATTC, 2014[11]).

For diploid cultures, there is a correlation between PDL and passage number which in turn depends on the growth surface/volume area and the initial seeding density. Some test facilities prefer to use for tracking and reporting cellular age PDL and not passage number especially for cells where (1) growth may vary significantly between donors and between preparations, (2) to correlate directly PDL number with replicative senescence which can be linked to specific phenotypic characteristics (e.g., loss of potency in mesenchymal stromal cells), (3) to correlated PDL numbers directly with genomic instability and (4) to use PDL as a standard for the new cell preparations to compare and analyse different studies[9].

Passage number refers to the number of times the cell line has been re-plated (adherent cultures) or re-seeded (suspension cultures). For adherent or suspension cultures, each reseeding (dilution) of the cells increases the passage number by one. Cultures should be subcultured while still in the log phase, i.e. before reaching confluence/saturation.

Each test facility should have SOPs in place, where details are provided not only about how to thaw, handle, count, maintain, bank and store their cell cultures, as well as for screening for contamination, but also to univocally assign progressive passage numbers and how to determine the cell stock viability.

The frequency of passaging (transfer between flasks with or without cell dilution) depends on the growth rate of the cell culture (adherent or in suspension) and the seeding density at passage (split ratio). Many dividing primary human cell cultures have a split ratio of one in two (1:2), while continuous cell lines have much higher splitting rates, e.g., atypical split ratio is between 1:3 and 1:8. The cells can take much longer to resume exponential growth if they are split at higher dilution ratios. It should be remembered that passaging will initially result in a loss of cells. The proportion of cells lost is variable and depends on the type of cell culture, the expertise of the operator and the plating efficiency (the proportion of cells that reattach) in the case of adherent cultures.

Some cell lines require a fixed seeding density and subculturing scheme where counting the number of cells is required (Wilson et al., 2015[13]) A more specific example is 3T3-Swiss or NIH/3T3 cell lines which were established by the same subculturing scheme (3T3 is a designation that stands for passaging the cells 3 times/week at 1:3) which is important to maintain the cell characteristics[10]. To improve consistency across experiments, all routine cell culture should utilise a fixed and pre-determined seeding density as estimations of cell confluency are prone to error and contribute to variability in baseline cell physiology. Most commonly cell counting is performed using the Bürker Türk or Neubauer counting chambers. When automated cell counters are used, their correct functioning would need to be demonstrated (Cadena-Herrera et al., 2015[14]; Gunetti et al., 2012[15]; Phelan and Lawler, 2001[16]). Cell viability, using Trypan Blue stain or other nuclear counterstains (Annex I), is commonly performed so as to count only viable cells for accurate seeding density calculation.

Different cell lines have different growth rates which may depend on several environmental factors. Many diverse culturing techniques have been used to fully reproduce the various environments test systems normally encounter during development. Most of the work to date has been performed on solid plastic supports including high-throughput plastic supports. A plastic growth support has several limitations in its representation of the *in vivo* environment. As plastic is an impermeable smooth two-

dimensional surface, it forces the cells to exchange their gas and nutrients exclusively through the top side of the cultured cells while *in vivo* cell are exposed from several directions to factors from the blood, other cells, soluble factors, and liquid-air interfaces. Growth of cells in more physiological conditions such as air-liquid interface set-ups (Pezzulo *et al*., 2011) or on microporous membranes (Klein et al., 2013[17]), or by using a variety of biomaterial coated surfaces for specific cell attachment, propagation, differentiation, and migration requirements (Chai and Leong, 2007[18]; Tallawi et al., 2015[19]) has many advantages and may be applied when examining aspects such as:

- Permeability and transport of macromolecules, ions, hormones, growth factors, and other biologically relevant molecules
- Cell polarity e.g., sorting and targeting of macromolecules; and polarized distribution of ion channels, enzymes, transport proteins, receptors and lipids
- Endocytosis
- Tumour invasion and metastasis
- Chemotaxis and other cell motility studies including angiogenesis, phagocytosis
- Co-culture effects, including interactions of feeder layers with stem cell cultures and cell-to-cell/matrix interactions
- Microbial pathogenesis e.g., test item effects on microbial receptors
- *In vitro* fertilisation including small molecule transport studies

Another advantage of cells grown on porous membrane substrates is their ability to provide a surface that better mimics a three-dimensional *in vivo* setting important for tissue remodelling (e.g., wound healing). Porous membranes allow multidirectional exposure to nutrients and waste products. Membrane separation of dual chambers allows for the co-culture of cells of different origin, and is used to study how cells interact through indirect signalling or through providing a conditioned niche for the proper growth and differentiation of cell types. Permeable supports also permit culture of polarised cells (Sheridan et al., 2008[20]).

If required for the particular test system, justification of the method for differentiation should be described in the *in vitro* method, since potential of differentiation and the method used to induce differentiation will vary depending on the type of cells, and should include justification of the process in which the method was determined.

5.4.1. Influence of the quality of the feeder layer

The growth of stem cells in culture requires certain nutrients that support the cells in an undifferentiated state. In this case a feeder cell layer is often used. One consideration in minimising variability of *in vitro* testing using stem cells is to ensure standardised methodology in deriving, culturing, and inactivating feeder cells. There are many kinds of cell lines used as feeder cell layer. Fibroblasts like mouse embryonic fibroblast cells and mouse embryo derived thioguanine and ouabain-resistant cell lines are commonly used to establish and culture embryonic stem cells (ESCs). Cell lines derived from umbilical cord blood cells or adult bone marrow cells have been used as ESC feeder cell layers. The influence of the quality and type of feeder layer can affect some pluripotency marker genes and proteins in ESC cultures (Healy and Ruban, 2015[21]).

5.5. Cryopreservation and thawing

Cryostorage systems should ensure the long term preservation of the stored test system. For cryopreserved cell cultures, the viability of mammalian cells is progressively lost

within months at -80˚C, thus, long term storage below the glass transition point of water (-136˚C) is recommended. While true for mammalian cells, this is not the case for bacteria or yeast.

Improved technologies that allow cryopreservation of *in vitro* cell and tissue cultures at different stages of differentiation and their long-term storage has introduced new or more standardised *in vitro* test systems into the pipeline of potential *in vitro* methods to be used in human safety assessment. Controlled-rate and slow freezing, also known as slow programmable freezing has been used all over the world for freezing down cell and tissue cultures. New methods are constantly being investigated due to the inherent toxicity of many cryoprotectants. Depending on the type of cell culture, using dimethyl sulfoxide (DMSO) may not always be preferable, as DMSO shows relatively strong cytotoxicity to some cells types and affects differentiation of iPS and ES cells (Katkov et al., 2006[22]).

As described in GCCP Principle 1, 'Establishment and maintenance of a sufficient understanding of the *in vitro* system and of the relevant factors which could affect it' (Coecke et al., 2005[5]), it is essential to prepare preserved banks of cells intended for use, to assure that reliable stocks can be obtained for testing which are at a consistent passage level from the original 'seed stock'. This is in order to avoid the effects of changes or cross-contamination which may occur if cell lines are maintained indefinitely. Standard cryopreservation methods using DMSO (10%) and serum (20%) as cryoprotectants, combined with a slow cooling rate (e.g., -1°C/min) and standardised cell numbers per vial will usually be successful for most cells. However, it is necessary to check the viability of preserved stocks in case of freezing failure and also to try to assure consistency between individual vials in a cell bank regarding cell number, viability and desired functionality. Viability measurements made immediately post-thaw can give misleadingly high values as many cells can be lost during the 24h recovery phase post thawing.

When stored in liquid nitrogen, storage in the vapour phase (Table 5.1. Comparison of ultra-low cryostorage methods for cells) is generally advised for all cells and necessary for potentially infectious cells and tissues. This eliminates the chances of transfer of pathogenic material between vials which can occur in the liquid phase of nitrogen (Coecke et al., 2005[5]). It is also considered safer as liquid nitrogen can enter storage vials if they are stored in the liquid phase which may cause them to explode upon thawing. However, to accommodate storage in the vapour phase, the amount of liquid nitrogen needs to be reduced, which will require more frequent topping up of the liquid nitrogen so as to maintain the correct storage temperature. If vials need to be stored in the liquid phase, protection wrapping may be considered.

Cryostorage requires temperature and/or liquid nitrogen level monitoring to ensure that the test system stocks are at optimal storage temperature. Cryostorage vessels can be fitted with alarms and data loggers and liquid nitrogen levels recorded at regular intervals (e.g., weekly). Appropriate safety protection (e.g., wearing of safety glasses, gloves etc.) should always be used when working with liquid nitrogen as contact with the skin or eyes may cause serious freezing (frostbite) or other injury.

Table 5.1. Comparison of ultra-low cryostorage methods for cells

Method	Advantages	Disadvantages
Electric Freezer (-130°C or lower)	Ease of Maintenance Steady temperature Low running costs	Requires CO_2, liquid N_2 or electrical backup Mechanically complex Storage temperatures high relative to liquid nitrogen
Liquid Phase Nitrogen	Steady ultra-low (-196°C) temperature Simplicity and mechanical reliability	Requires regular supply of liquid nitrogen High running costs Risk of cross-contamination via the liquid nitrogen
Vapour Phase Nitrogen	No risk of cross-contamination from liquid nitrogen Low temperatures achieved Simplicity and mechanical reliability	Requires regular supply of liquid nitrogen High running costs Temperature fluctuations between -135°C and -190°C

Source: (ECACC , 2010[23])

Storing valuable test system stocks in more than one cryostorage location is recommended for security/backup purposes and off-site storage may also need to be considered in disaster recovery plans for the facility.

A number of factors may affect the viability of cells on thawing, including the composition of the freeze medium, the growth phase of the culture, the stage of the cell in the cell cycle, and the number and concentration of cells within the freezing solution (ATTC, 2014[11]). Another issue to take into consideration when using thawed cells is the possibility that the cells are stressed directly after thawing, which appears to involve apoptosis (Baust, Van Buskirk and Baust, 2002[24]). Most cells begin to recover after 24 hours and enter the log (exponential) growth phase soon afterwards. It is therefore necessary to remove DMSO and any dead cells as they might affect the seeding density calculation. It is also recommended not to use them straight away, but to passage them at least twice, so as to allow the cells to re-establish their normal cell cycle.

Optimum conditions should be defined in the *in vitro* method SOP(s) during the development phase. When the test system is sourced from a commercial supplier (Section 5.2) extensive documentation is usually provided including detailed information for handling the cells, including cryopreservation and thawing information. Batch-specific information such as the number of cells per vial, the recommended split or subcultivation ratio, and the passage number and/or population doubling level (PDL) when known are also provided.

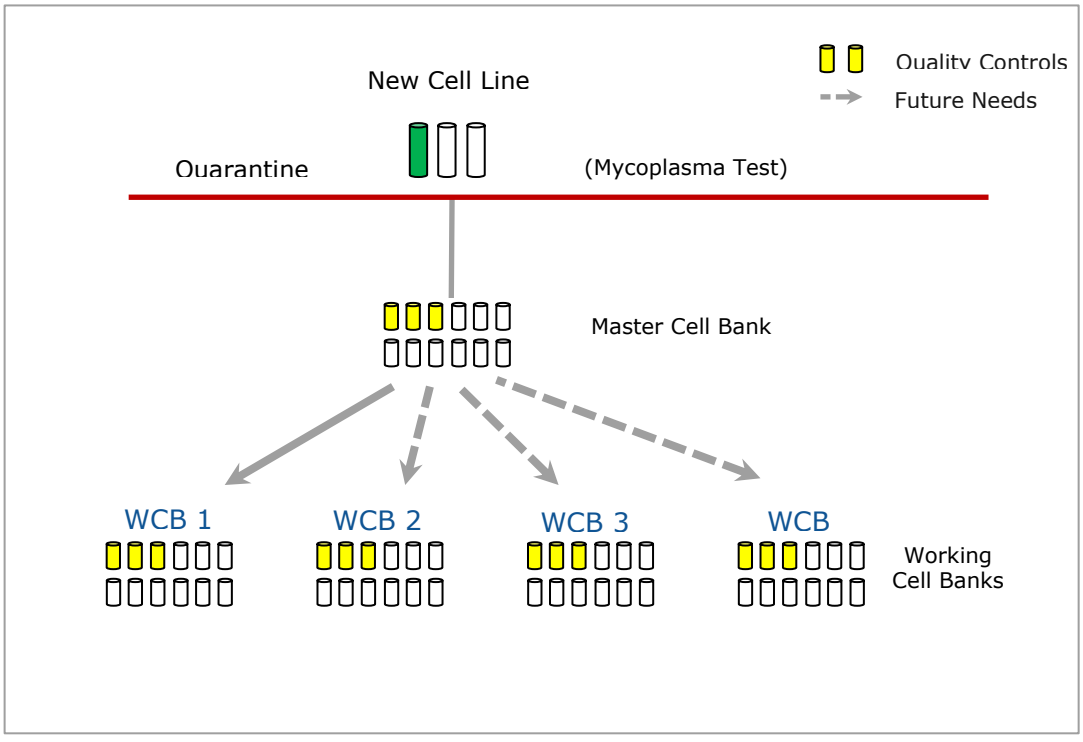
5.5.1. Cell Banking

Maintaining a cell line in continuous or extended culture is considered bad practice as there may be a higher risk of microbial contamination and/or cross contamination with other cell lines, a loss of characteristics of interest, genetic drift particularly in cells known to have an unstable karyotype or loss due to exceeding finite life-span[11]. It is therefore important to avoid subjecting cell lines to variable culture and passage conditions, and to establish cryopreserved stocks of early passage cells (Coecke et al., 2005[5]).

New cell lines should be quarantined (Section 3.5) until their origin has been authenticated and they are shown to be free of microorganisms (Geraghty et al., 2014[25]).

A two-tiered cell banking system consisting of a Master Cell Bank (MCB) and a Working Cell Bank (WCB) is recommended. Cells from the new cell line are placed in culture and harvested when they are at their maximum growth rate or almost confluent. These are then frozen to create a master cell bank, usually consisting of 10 to 20 vials of 1 ml, each usually containing 1–5 × 106 cells (Geraghty et al., 2014[25]). The MCB is not for distribution and should be protected from unintended use, however new working banks may be created from the original master bank when required (Figure 5.2). From this MCB, a single vial is thawed and cultured until there are enough cells to produce a working cell bank. The working cell bank should contain sufficient ampoules to cover the proposed experimental period plus sufficient ampoules for contingencies and distribution (UKCCCR, 2000[26]).

Figure 5.2. Flexible two tiered approach to cell banking



Quality controls procedures defined in SOPs should include checks on viability, mycoplasma and other microbial contaminants, cell line identity and any other relevant cell line characteristics (Coecke et al., 2005[5]), and should be applied systematically to the working cell banks. Quality Control (QC) tests for the absence of bacteria and fungi, and testing for mycoplasma should only be performed following a period of antibiotic-free culture. For primary cells, the state of cell differentiation should also be carefully observed during banking. Different passage of primary cells with different differentiation status will greatly influence results.

5.6. Cell line identity and genetic aberrations

Genetic, phenotypic and immunological markers are useful in establishing the identity of the cell(s). Genetic stability testing (also known as cell line stability) is a key component

in characterising cell banks and is especially critical in maintaining quality of mammalian cell cultures. For an engineered cell line, the inserted gene of interest should remain intact and at the same copy number, and be expressed. Furthermore, there should be traceability to the original provider of the cell culture and the related documentation. However, a frequent problem in the use of cell culture is the use of cells which have become cross-contaminated, misidentified (see International Cell Line Authentication Committee (ICLAC) database of cross contaminated or misidentified cell lines[12]), mixed-up, or present genomic instability (Allen et al., 2016[27]); (Frattini et al., 2015[28]); (Fusenig et al., 2017[29]); (Kleensang et al., 2016[30]); (Vogel, 2010[31]). This is not always detectable by cell morphology and/or culture characteristics. An example of a mistake from the past in an OECD test guideline method (OECD, 2016[32]) is BG1Luc4E2 cells which have been renamed VM7Luc4E2 cells. The reason being recent DNA analysis revealed that the original cell line used to generate the BG1Luc4E2 cells were not human ovarian carcinoma (BG1) cells but a variant of human breast cancer (MCF7) cells[13].

There are different genomic techniques for human and non-human cell line authentication (Table 5.2). Cell line authentication is an example of the kind of data that add confidence to the results of a scientific study. The lack of reporting of cell line authentication data reflects a broader failure to appreciate the need for more complete reporting of experimental details that qualify data and provide confidence in the scientific results (Almeida, Cole and Plant, 2016[33]); (Marx, 2014[34]).

Table 5.2. Current status of SNP, STR, and DNA barcode technologies as standard methods for assessing the identity of cell lines from different species

Species	Assays	Consensus Standard Method	Commercially Available Kit	Commercial Service	Comparative Data
Human	STR	ASN-0002	Yes	Yes	ATCC, DSMZ, JCRB, NCBI**
	SNP	No	Yes	Yes	(Liang-Chu et al., 2015[35]) (Yu et al., 2015[36]) NCBI
Mouse	STR*	No	No	Yes	Unpublished
	SNP	No	Yes	Yes	(Didion et al., 2014[37])
African green monkey	STR*	No	No	No	None
Chinese hamster ovary	STR*	No	No	No	None
Rat	STR*	No	No	No	None
Species-level identification	CO1 DNA barcode	ASN-0003	Yes	Yes	Barcode of Life Data System, NCBI**
	Species-specific primers	No	No	Yes	None needed

Note: These methods are currently the most developed for this application. There are extensive data on human cell lines, but while there are some kits and services for some nonhuman species, there is little available data for nonhuman species, except for DNA barcoding, which only distinguishes cell lines on the basis of species, not individuals.
* STR markers have been identified (Almeida, Hill and Cole, 2013[38]); (Almeida, Hill and Cole, 2011[39]). Markers for rat and Chinese hamster ovary cells are still under development by NIST.
** These sources contain a significant amount of data from multiple sources.
Source: (Almeida, Cole and Plant, 2016[33])

Establishing an early stock (or retention of a sample of original tissue) which is DNA fingerprinted will provide an important reference for future cell banks and for other centres. Short Tandem Repeat (STR) profiling is typically applied and has considerable background qualification for use in human samples (ISCBI, 2009[40]). STR profiling has been the subject of a comprehensive and definitive standard, ASN-0002[14], and can be performed in most laboratories that have the capabilities to execute molecular techniques. It is an easy, low cost and reliable method for the authentication of human cell lines. For non-human samples, Single Nucleotide Polymorphism analysis (aSNP), STR profiling, and DNA barcode technologies are available as methods for assessing the identity of cell lines from different species (Ono et al., 2007[41]). The field of genetic analysis is progressing rapidly and interested parties should maintain knowledge of current best scientific practice in this area as next generation sequencing begins to become a routine tool. Doing so problems can be avoided early in the process and not jeopardise the cell lines used for regulatory purposes as has happened for the Bhas 42 cell line in a cell transformation *in vitro* method (OECD, 2016[42]) where issues related to misidentification arose at a late stage.

Genetic instability is inherent in cell cultures and it is wise to minimise the number of passages over which cells are maintained (typically 15-20). Although passage number alone is not a reliable parameter to ensure good cell functioning, it is good practice to define a limit for the maximum number of passages, possibly in combination with defined performance characteristics. At that limit, new cultures should be restarted from a working cell bank. The use of cells at higher passage numbers must be justified and their integrity and functionality demonstrated. Where cells are known to be extremely unstable, some form of genetic testing, such as karyology or molecular analysis like aSNP or Comparative Genomic Hybridisation (aCGH) may need to be performed. In particular, this applies to recombinant cell lines including those maintained with antibiotic selection. Recombinant cell lines should be maintained in parallel with matched cells that were generated with the empty vector alone and were simultaneously subjected to antibiotic selection. Such matched cells will be a more suitable control than non-selected cells that do not express the same resistance marker as their modified counterparts.

There are special issues for stem cells. Stem cell lines may contain a mixture of diploid and aneuploid cells, which may be unavoidable, but genetic testing (see above) can be used to screen for progressive change (e.g., between master and working cell banks) which could impact on the suitability of the cell culture. Human induced Pluripotent Stem Cell (hiPSC) lines should also be tested for absence of ectopic expression of reprogramming genes and where produced by non-integrating vectors, for elimination of the vector. iPS/ESCs also need to be evaluated by their genotypes and differentiation capability by embryoid body formation, direct differentiation method and teratoma formation assays.

Acceptable intervals for periodic testing to confirm the genetic, phenotypic and immunological stability of the cell culture are highly case-dependent (Blázquez-Prunera et al., 2017[43]) (Daily et al., 2017[44]; Meza-Zepeda et al., 2008[45]). Therefore, this aspect should be included in the specific test system SOP(s).

5.7. Contaminants screening: sterility, mycoplasma, virus

Standard sterility tests are widely available[15] [16] and may be used for cell stocks and cultures; however, it is important to bear in mind that these tests are usually based on inoculation of broth cultures which may not support the growth of all contaminating

microorganisms. Alternative molecular methods such as identification by Polymerase Chain Reaction (PCR) and DNA sequencing of ribosomal RNA may be used.

Viruses may arise as contaminants of cell cultures via the original donor used to produce the cell line or feeder cells and other biological reagents used in cell culture. They may cause cytopathic effects, in which case the culture should be discarded, or they may have no effect and become diluted out when fresh uncontaminated reagents are used. In certain cases they may establish persistent infections, although this is believed to be rare. Whatever the outcome, their presence and influence on cell biology may be significant as amongst other effects they may modify transcription factor networks and alter the cells' biology. To assure laboratory worker safety, some organisations require testing of all human cell lines for serious human pathogens such as Human Immunodeficiency Virus (HIV) and Hepatitis B and C or evidence that the donors did not have these pathogens. However, such testing clearly does not cover more common human infections, and human pathogens may also be carried by cells from other species. Cell line testing may be initiated if there are special hazards associated with the work or with the cells. Workers should always follow local rules for performing cell culture work, maintain their competence in aseptic processing, as well as carry out regular and careful inspection of cells for any unusual effects or morphologies that might indicate infection. A robust testing regime for contamination should include procedures for managing positive results, whether to immediately discard or quarantine the affected cells until a means of action can be decided along with the detection of the root cause by supplementary testing (Stacey, 2011[46]).

It is crucial to routinely test cell cultures for the presence of mycoplasma. A range of test techniques are available and it is advised for critical tests to use methods which detect cultivable as well as non-cultivable mycoplasma species, e.g. PCR-based methods (Table Table 5.3). It is important to know what aspect of contamination the test is designed to detect, how well the test performs, its specificity (i.e., what strains of mycoplasma its detects and any likely causes of false positive reactions) and for detectable contamination, what level of sensitivity is achievable under the prescribed sampling and test conditions. Selection of test methods should be based on evaluation of the potential specificity and sensitivity of detection and the likelihood of inhibition of a positive result.

EMA has provided a general chapter on mycoplasma testing of cells which should be consulted (EMA, 2013[47]). All aspects of the test sample which are likely to influence the strains which may be isolated and any conditions which may affect detection such as inhibitory substances, should be evaluated before selecting a particular technique. Even where alternative detection kits are based on the same basic methodology, their specificity and sensitivity may vary considerably and even the same methods used in different laboratories may be influenced by local differences in raw materials, test conditions and the way the test is performed. Accordingly, any test method used should be subjected to the general evaluation indicated above and performance of testing should be accompanied by the inclusion of appropriate controls as below:

1. positive controls (including a reference sample close to the limit of detection),
2. negative controls to exclude false positives from reagents and test conditions and
3. positive controls spiked into test samples (or other approaches to control for sample inhibition) of positive test results.

All such testing should be performed only by a person trained and competent in the test. Records of performance, including positives and negative results, control performance

and any equivocal or anomalous results, and any trends in quantitative results for test and control samples (where applicable) should be kept, so as to enable ongoing evaluation.

Table 5.3. Mycoplasma detection methods, their sensitivity, and advantages and disadvantages

Method	Sensitivity	Advantages	Disadvantages
Indirect DNA stain (e.g., Hoechst 33258) with indicator cells (e.g., 3T3)	High	Easy to interpret because contamination amplified	Indirect and thus more time-consuming
Broth and agar culture	High	Sensitive	Slow and may require expert interpretation
PCR	High	Rapid	Requires optimisation
Nested PCR	High	Rapid	More sensitive than direct PCR, but more likely to give false positives
Enzyme-Linked Immunosorbent Assay (ELISA)	Moderate	Rapid	Limited range of species detected
Autoradiography	Moderate	Rapid	Can be difficult to interpret if contamination is at low level
Immunostaining	Moderate	Rapid	Can be difficult to interpret if contamination is at low level
Direct DNA stain (e.g., Hoechst 33258)	Low	Rapid, cheap	Can be difficult to interpret

Source: (Young et al., 2010[48])

5.8. Biomarkers and functional tests to confirm the required cell function state

It is important to recognise that cell quality can vary during passaging, and in particular the time point in the growth curve at which cells are harvested (ideally in the log or exponential phase) may affect performance (Section 5.4). Accordingly, each culture used to set up an *in vitro* method should be subject to a key control regime measuring or indicating functionality. Because the crucial function of the test system to be measured may be dependent on the last step in a long sequence of events (e.g., gene activation and gene-transcript-protein-reaction) it is of importance to ensure that a selected biomarker or a test is directly related to the crucial function to be measured. Acceptance criteria should be defined for functional tests and biomarkers that indicate the correct cell state. These may for example include: neuronal activity, competency of biochemical transformation, response to reference bioactive compounds when using metabolically competent cells, response to reference items in the particular *in vitro* method the cells are to be used for etc. In this way, each culture can be controlled, and consistency in *in vitro* methods is supported. For example, expression of self-renewal genes (e.g., Oct4, Nanog, Sox2) in stem cell cultures is crucial to the functionality of the cell population (further examples for stem cells are laid out in (Pistollato et al., 2014[49]; Stacey et al., 2016[6]). Additionally, key markers which are associated with poor performance may be identified for future improvement.

As stem cell-based (both hiPSC and hESC) *in vitro* models have and will be employed for regulatory use, not only key markers for cell state but also the maturation phase requires characterisation. For example, human pluripotent stem cell-derived cardiomyocytes have been shown to display morphological and functional properties typical of human foetal cardiomyocytes which may complicate their utilisation and interpretation of the obtained results (Robertson, Tran and George, 2013[50]; Snir et al., 2003[51]). Increased time in culture, electrical stimulation (Chan et al., 2013[52]) and 3D culture environment (Garzoni et al., 2009[53]; Schaaf et al., 2011[54]; Soares et al., 2012[55]; Valarmathi et al., 2010[56];

van Spreeuwel et al., 2014[57]) have been utilised in the production of more mature cardiomyocytes with adult-like properties. As such, the maturation phase of stem cells is deemed a critical quality parameter when the relevance of an *in vitro* test system is to be considered.

In a co-culture system, the use of stem cells provide the test system with multipotent differentiation capacity and can act as helper cells for ensuring homeostasis, metabolism, growth and recovery. Their inclusion in co-culture systems has shown benefits creating complex tissues, including orthopaedic soft tissues, bone, heart, blood vessels, lungs, kidneys, liver and nerves (Paschos et al., 2014[58]). In addition, it is necessary to evaluate both combination of biomarkers and cytometric analysis (e.g., flow cytometry or fluorescent microscopy) to check robustness of the stem cell culture.

Interactions within the same cell population (homotypic) and between different cell types (heterotypic or co-cultures) are essential for tissue development, repair, and homeostasis. Some cells cannot easily be mono-cultured *in vitro* or at least do not exhibit desired *in vivo* physiological behaviours, but the presence of another cell population may improve the culturing success or cell behaviour. Cell-cell interactions in co-cultures are strongly influenced by the extracellular environment, which is determined by the experimental set-up, which therefore needs to be given careful consideration (Goers, Freemont and Polizzi, 2014[59]). It is critical to identify biomarkers and functional tests to confirm the required co-culture system function state.

5.9. Metabolic activation

Metabolism is a bottleneck in *in vitro* toxicological method development since there is an inability of most mammalian *in vitro* cell and tissue cultures to predict the physiological effect of *in vivo* metabolism by the Phase I and Phase II biotransformation enzymes (Coecke et al., 2006[60]).

Currently no *in vitro* cell and tissue culture test system will mirror fully the complexity of *in vivo* metabolism, and the production of active metabolites may either not occur in non-metabolic competent test systems or be over or underestimated in metabolically-competent test systems. However, these considerations should not prevent the use of metabolic activation systems to mitigate this problem, provided the limitations and drawbacks are clearly understood and the results take into consideration these limitations.

The evolution of genotoxicity testing offers a good example of the use of metabolic activation mixtures to improve the physiological relevance of *in vitro* methods for genotoxicity testing[17]. In the case of the Ames test (OECD, 1997[61]) a metabolising system in the form of a cofactor-supplemented S9 fraction, containing microsomal and cytosolic fractions prepared from rat liver (usually) pre-treated with enzyme inducing agents such as Arochlor 1254 or a combination of phenobarbitone and ß-naphthoflavone, to induce metabolising enzymes, has been built into the method. In 1997 the OECD issued a detailed review paper (DRP) on the use of metabolising systems for *in vitro* testing of endocrine disruptors detailing different options how to produce the relevant metabolites of the test item under investigation when carrying out these types of tests. It is recommended that *in vitro* method developers take this aspect into consideration when designing *in vitro* method(s) (OECD, 2008[62]). Furthermore, there is a need for metabolically-active test systems both for toxicokinetics and toxicodynamics applications in regulatory testing (Chan et al., 2013[52]).

Possible strategies how to employ metabolic activation when designing *in vitro* methods remain a challenge even for well-established methods (Nesslany, 2017[63]). However, efforts are underway to introduce a metabolic component in OECD TG methods for the detection of chemicals with (anti)estrogenic potential (OECD, 2016[32]).

More and more integrated ways to predict the physiological effect of metabolism are being proposed in response to the open challenges for regulatory toxicology (Funk and Roth, 2016[64]; Wang et al., 2013[65]; Williams et al., 2013[66]).

Notes

1. See: http://wiki.toxbank.net/w/images/1/18/ToxBank_D4_6_final_10_04_13.pdf
2. GCCP Principle 3: Documentation of the information necessary to track the materials and methods used, to permit the repetition of the work, and to enable the target audience to understand and evaluate the work
3. See: https://hpscreg.eu/
4. See: http://bch.cbd.int/protocol/text/
5. See: http://www.iata.org/whatwedo/cargo/dgr/Documents/infectious-substance-classification-DGR56-en.pdf
6. See: http://www.iata.org/whatwedo/cargo/dgr/Pages/download.aspx
7. See: http://www.un3373.com/un3373-packaging/un3373/
8. See: https://www.atcc.org/en/Global/FAQs/B/C/Passage_number_vs_population_doubling_level_PDL-175.aspx
9. See: http://stemcellassays.com/2014/05/msc-pdl/
10. See: https://www.atcc.org/Global/FAQs/8/9/ATCC%20CCL92%20vs%20ATCC%20CRL1658-453.aspx
11. See: http://www.sigmaaldrich.com/technical-documents/protocols/biology/good-cell-banking.html
12. See: http://iclac.org/databases/cross-contaminations/
13. See: https://ntp.niehs.nih.gov/iccvam/methods/endocrine/bg1luc/bg1luc-vm7luc-june2016-508.pdf
14. See: http://webstore.ansi.org
15. See: http://www.who.int/medicines/publications/pharmacopoeia/TestForSterility-RevGenMethod_QAS11-413FINALMarch2012.pdf
16. See: http://medicaldesign.com/site-files/medicaldesign.com/files/archive/medicaldesign.com/Whitepapers/SterilityTestin_00000021071.pdf
17. See: https://www.oecd.org/chemicalsafety/testing/Genetic%20Toxicology%20Guidance%20Document%20Aug%2031%202015.pdf

References

Allen, M. et al. (2016), "Origin of the U87MG glioma cell line: Good news and bad news", *Science Translational Medicine*, Vol. 8/354, pp. 354re3-354re3, http://dx.doi.org/10.1126/scitranslmed.aaf6853. [27]

Almeida, J., K. Cole and A. Plant (2016), "Standards for Cell Line Authentication and Beyond", *PLOS Biology*, Vol. 14/6, p. e1002476, http://dx.doi.org/10.1371/journal.pbio.1002476. [33]

Almeida, J., C. Hill and K. Cole (2013), "Mouse cell line authentication", *Cytotechnology*, Vol. 66/1, pp. 133-147, http://dx.doi.org/10.1007/s10616-013-9545-7. [38]

Almeida, J., C. Hill and K. Cole (2011), "Authentication of African green monkey cell lines using human short tandem repeat markers", *BMC Biotechnology*, Vol. 11/1, p. 102, http://dx.doi.org/10.1186/1472-6750-11-102. [39]

Anderson, R. et al. (1998), "The Availability of Human Tissue for Biomedical Research: The Report and Recommendations of the ECVAM Workshop 32.", *Alternatives to laboratory animals*, pp. 763-777. [7]

ATTC (2014), *Animal Cell Culture Guide*, ATTC. [11]

Baust, J., R. Van Buskirk and J. Baust (2002), "Gene Activation of the Apoptotic Caspase Cascade Following Cryogenic Storage", *Cell Preservation Technology*, Vol. 1/1, pp. 63-80, http://dx.doi.org/10.1089/15383440260073301. [24]

Bielecka, A. and A. Mohammadi (2014), "State-of-the-Art in Biosafety and Biosecurity in European Countries", *Archivum Immunologiae et Therapiae Experimentalis*, Vol. 62/3, pp. 169-178, http://dx.doi.org/10.1007/s00005-014-0290-1. [9]

Blázquez-Prunera, A. et al. (2017), "Human mesenchymal stem cells maintain their phenotype, multipotentiality, and genetic stability when cultured using a defined xeno-free human plasma fraction", *Stem Cell Research & Therapy*, Vol. 8/1, http://dx.doi.org/10.1186/s13287-017-0552-z. [43]

Cadena-Herrera, D. et al. (2015), "Validation of three viable-cell counting methods: Manual, semi-automated, and automated", *Biotechnology Reports*, Vol. 7, pp. 9-16, http://dx.doi.org/10.1016/j.btre.2015.04.004. [14]

Chai, C. and K. Leong (2007), "Biomaterials Approach to Expand and Direct Differentiation of Stem Cells", *Molecular Therapy*, Vol. 15/3, pp. 467-480, http://dx.doi.org/10.1038/sj.mt.6300084. [18]

Chan, Y. et al. (2013), "Electrical Stimulation Promotes Maturation of Cardiomyocytes Derived from Human Embryonic Stem Cells", *Journal of Cardiovascular Translational Research*, Vol. 6/6, pp. 989-999, http://dx.doi.org/10.1007/s12265-013-9510-z. [52]

Coecke, S. et al. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33/3, pp. 261-287. [5]

Coecke, S. et al. (2006), "Metabolism: a bottleneck in in vitro toxicological test development. The report and recommendations of ECVAM workshop 54.", *Alternatives to laboratory animals*. [60]

Daily, K. et al. (2017), "Molecular, phenotypic, and sample-associated data to describe pluripotent stem cell lines and derivatives", *Scientific Data*, Vol. 4, p. 170030, http://dx.doi.org/10.1038/sdata.2017.30. [44]

Didion, J. et al. (2014), "SNP array profiling of mouse cell lines identifies their strains of origin and reveals cross-contamination and widespread aneuploidy", *BMC Genomics*, Vol. 15/1, p. 847, http://dx.doi.org/10.1186/1471-2164-15-847. [37]

ECACC (2010), *Fundamentals Techniques in Cell Culture*, European Collection of Authenticated Cell Cultures, United Kingdom. [23]

EMA (2013), *VICH GL34: Biologicals: testing for the detection of Mycoplasma contamination*, European Medicines Agency. [47]

EU (2009), *Directive 2009/41/EC of the European Parliament and of the Council*. [10]

Frattini, A. et al. (2015), "High variability of genomic instability and gene expression profiling in different HeLa clones", *Scientific Reports*, Vol. 5/1, http://dx.doi.org/10.1038/srep15377. [28]

Funk, C. and A. Roth (2016), "Current limitations and future opportunities for prediction of DILI from in vitro", *Archives of Toxicology*, Vol. 91/1, pp. 131-142, http://dx.doi.org/10.1007/s00204-016-1874-9. [64]

Fusenig, N. et al. (2017), "The need for a worldwide consensus for cell line authentication: Experience implementing a mandatory requirement at the International Journal of Cancer", *PLOS Biology*, Vol. 15/4, p. e2001438, http://dx.doi.org/10.1371/journal.pbio.2001438. [29]

Garzoni, L. et al. (2009), "Dissecting coronary angiogenesis: 3D co-culture of cardiomyocytes with endothelial or mesenchymal cells", *Experimental Cell Research*, Vol. 315/19, pp. 3406-3418, http://dx.doi.org/10.1016/j.yexcr.2009.09.016. [53]

Geraghty, R. et al. (2014), "Guidelines for the use of cell lines in biomedical research", *British Journal of Cancer*, Vol. 111/6, pp. 1021-1046, http://dx.doi.org/10.1038/bjc.2014.166. [25]

Goers, L., P. Freemont and K. Polizzi (2014), "Co-culture systems and technologies: taking synthetic biology to the next level", *Journal of The Royal Society Interface*, Vol. 11/96, pp. 20140065-20140065, http://dx.doi.org/10.1098/rsif.2014.0065. [59]

Gunetti, M. et al. (2012), "Validation of analytical methods in GMP: the disposable Fast Read 102® device, an alternative practical approach for cell counting", *Journal of Translational Medicine*, Vol. 10/1, p. 112, http://dx.doi.org/10.1186/1479-5876-10-112. [15]

Healy, L. and L. Ruban (2015), *Atlas of Human Pluripotent Stem Cells in Culture*, Springer US, Boston, MA, http://dx.doi.org/10.1007/978-1-4899-7507-2. [21]

ISCBI (2009), "Consensus Guidance for Banking and Supply of Human Embryonic Stem Cell Lines for Research Purposes", *Stem Cell Reviews and Reports*, Vol. 5/4, pp. 301-314, http://dx.doi.org/10.1007/s12015-009-9085-x. [40]

Katkov, I. et al. (2006), "Cryopreservation by slow cooling with DMSO diminished production of Oct-4 pluripotency marker in human embryonic stem cells", *Cryobiology*, Vol. 53/2, pp. 194-205, http://dx.doi.org/10.1016/j.cryobiol.2006.05.005. [22]

Kleensang, A. et al. (2016), "Genetic variability in a frozen batch of MCF-7 cells invisible in routine authentication affecting cell function", *Scientific Reports*, Vol. 6/1, http://dx.doi.org/10.1038/srep28994. [30]

Klein, S. et al. (2013), "An improved 3D tetraculture system mimicking the cellular organisation at the alveolar barrier to study the potential toxic effects of particles on the lung", *Particle and Fibre Toxicology*, Vol. 10/1, p. 31, http://dx.doi.org/10.1186/1743-8977-10-31. [17]

Calafell, F. (ed.) (2015), "Human Biosample Authentication Using the High-Throughput, Cost-Effective SNPtraceTM System", *PLOS ONE*, Vol. 10/2, p. e0116218, http://dx.doi.org/10.1371/journal.pone.0116218. [35]

Lorge, E. et al. (2016), "Standardized cell sources and recommendations for good cell culture practices in genotoxicity testing", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 809, pp. 1-15, http://dx.doi.org/10.1016/j.mrgentox.2016.08.001. [4]

Marx, V. (2014), "Cell-line authentication demystified", *Nature Methods*, Vol. 11/5, pp. 483-488, http://dx.doi.org/10.1038/nmeth.2932. [34]

Meza-Zepeda, L. et al. (2008), "High-resolution analysis of genetic stability of human adipose tissue stem cells cultured to senescence", *Journal of Cellular and Molecular Medicine*, Vol. 12/2, pp. 553-563, http://dx.doi.org/10.1111/j.1582-4934.2007.00146.x. [45]

Nesslany, F. (2017), "The current limitations of in vitro genotoxicity testing and their relevance to the in vivo situation", *Food and Chemical Toxicology*, Vol. 106, pp. 609-615, http://dx.doi.org/10.1016/j.fct.2016.08.035. [63]

OECD (2016), *Guidance Document on the In Vitro Bhas 42 Cell Transformation Assay*, OECD Publishing, Paris, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2016)1&doclanguage=en. [42]

OECD (2016), *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264265295-en. [32]

OECD (2008), *Detailed Review Paper on the Use of Metabolising Systems for In Vitro Testing of Endocrine Disruptors*, OECD Publishing, Paris. [62]

OECD (1997), *Test No. 471: Bacterial Reverse Mutation Test*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264071247-en. [61]

Ono, K. et al. (2007), "Species identification of animal cells by nested PCR targeted to mitochondrial DNA", *In Vitro Cellular & Developmental Biology - Animal*, Vol. 43/5-6, pp. 168-175, http://dx.doi.org/10.1007/s11626-007-9033-5. [41]

Pamies, D. (2016), "Good Cell Culture Practice for stem cells and stem-cell-derived models", *ALTEX*, http://dx.doi.org/10.14573/altex.1607121. [3]

Paschos, N. et al. (2014), "Advances in tissue engineering through stem cell-based co-culture", *Journal of Tissue Engineering and Regenerative Medicine*, Vol. 9/5, pp. 488-503, http://dx.doi.org/10.1002/term.1870. [58]

Phelan, M. and G. Lawler (2001), "Cell Counting", *Current Protocols in Cytometry*, Vol. 00/1, pp. A.3A.1-A.3A.4, http://dx.doi.org/10.1002/0471142956.cya03as00. [16]

Pistollato, F. et al. (2014), "Development of a pluripotent stem cell derived neuronal model to identify chemically induced pathway perturbations in relation to neurotoxicity: Effects of CREB pathway inhibition", *Toxicology and Applied Pharmacology*, Vol. 280/2, pp. 378-388, http://dx.doi.org/10.1016/j.taap.2014.08.007. [49]

Riss, T. and R. Moravec (2004), "Use of Multiple Assay Endpoints to Investigate the Effects of Incubation Time, Dose of Toxin, and Plating Density in Cell-Based Cytotoxicity Assays", *ASSAY and Drug Development Technologies*, Vol. 2/1, pp. 51-62, http://dx.doi.org/10.1089/154065804322966315. [12]

Robertson, C., D. Tran and S. George (2013), "Concise Review: Maturation Phases of Human Pluripotent Stem Cell-Derived Cardiomyocytes", *STEM CELLS*, Vol. 31/5, pp. 829-837, http://dx.doi.org/10.1002/stem.1331. [50]

Schaaf, S. et al. (2011), "Human Engineered Heart Tissue as a Versatile Tool in Basic Research and Preclinical Toxicology", *PLoS ONE*, Vol. 6/10, p. e26397, http://dx.doi.org/10.1371/journal.pone.0026397. [54]

Sheridan, S. et al. (2008), "Microporous Membrane Growth Substrates for Embryonic Stem Cell Culture and Differentiation", in *Methods in Cell Biology, Stem Cell Culture*, Elsevier, http://dx.doi.org/10.1016/s0091-679x(08)00003-4. [20]

Snir, M. et al. (2003), "Assessment of the ultrastructural and proliferative properties of human embryonic stem cell-derived cardiomyocytes", *American Journal of Physiology-Heart and Circulatory Physiology*, Vol. 285/6, pp. H2355-H2363, http://dx.doi.org/10.1152/ajpheart.00020.2003. [51]

Soares, C. et al. (2012), "2D and 3D-Organized Cardiac Cells Shows Differences in Cellular Morphology, Adhesion Junctions, Presence of Myofibrils and Protein Expression", *PLoS ONE*, Vol. 7/5, p. e38147, http://dx.doi.org/10.1371/journal.pone.0038147. [55]

Soldatow, V. et al. (2013), "In vitro models for liver toxicity testing", *Toxicol. Res.*, Vol. 2/1, pp. 23-39, http://dx.doi.org/10.1039/c2tx20051a. [2]

Stacey, G. (2011), "Cell Culture Contamination", in *Methods in Molecular Biology, Cancer Cell Culture*, Humana Press, Totowa, NJ, http://dx.doi.org/10.1007/978-1-61779-080-5_7. [46]

Stacey, G. et al. (2016), "Ensuring the Quality of Stem Cell-Derived In Vitro Models for Toxicity Testing", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_11. [6]

Stacey, G. and T. Hartung (2006), "Availability, Standardization and Safety of Human Cells and Tissues for Drug Screening and Testing", in *Drug Testing in vitro*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany, http://dx.doi.org/10.1002/9783527609611.ch9. [8]

Tallawi, M. et al. (2015), "Strategies for the chemical and biological functionalization of scaffolds for cardiac tissue engineering: a review", *Journal of The Royal Society Interface*, Vol. 12/108, p. 20150254, http://dx.doi.org/10.1098/rsif.2015.0254. [19]

UKCCCR (2000), "UKCCCR Guidelines for the Use of Cell Lines in Cancer Research", *British Journal of Cancer*, Vol. 82/9, pp. 1495-1509, http://dx.doi.org/10.1054/bjoc.1999.1169. [26]

Valarmathi, M. et al. (2010), "A 3-D cardiac muscle construct for exploring adult marrow stem cell based myocardial regeneration", *Biomaterials*, Vol. 31/12, pp. 3185-3200, http://dx.doi.org/10.1016/j.biomaterials.2010.01.041. [56]

van Spreeuwel, A. et al. (2014), "The influence of matrix (an)isotropy on cardiomyocyte contraction in engineered cardiac microtissues", *Integr. Biol.*, Vol. 6/4, pp. 422-429, http://dx.doi.org/10.1039/c3ib40219c. [57]

Vogel, G. (2010), "To Scientists' Dismay, Mixed-Up Cell Lines Strike Again", *Science*, Vol. 329/5995, pp. 1004-1004, http://dx.doi.org/10.1126/science.329.5995.1004. [31]

Wang, S. et al. (2013), "Towards an integratedin vitrostrategy for estrogenicity testing", *Journal of Applied Toxicology*, Vol. 34/9, pp. 1031-1040, http://dx.doi.org/10.1002/jat.2928. [65]

Watson, D., R. Hunziker and J. Wikswo (2017), "Fitting tissue chips and microphysiological systems into the grand scheme of medicine, biology, pharmacology, and toxicology", *Experimental Biology and Medicine*, Vol. 242/16, pp. 1559-1572, http://dx.doi.org/10.1177/1535370217732765. [1]

Williams, D. et al. (2013), "Novel in vitro and mathematical models for the prediction of chemical toxicity", *Toxicol. Res.*, Vol. 2/1, pp. 40-59, http://dx.doi.org/10.1039/c2tx20031g. [66]

Wilson, H. et al. (2015), "Exploring the effects of cell seeding density on the differentiation of human pluripotent stem cells to brain microvascular endothelial cells", *Fluids and Barriers of the CNS*, Vol. 12/1, http://dx.doi.org/10.1186/s12987-015-0007-9. [13]

Young, L. et al. (2010), "Detection of Mycoplasma in cell cultures", *Nature Protocols*, Vol. 5/5, pp. 929-934, http://dx.doi.org/10.1038/nprot.2010.43. [48]

Yu, M. et al. (2015), "A resource for cell line authentication, annotation and quality control", *Nature*, Vol. 520/7547, pp. 307-311, http://dx.doi.org/10.1038/nature14397. [36]

Chapter 6. Test and reference/control items

Key message: *The preparation and characterisation of test, reference and control items and their interaction with the in vitro environment should be well understood so as to ensure the acquisition of reliable and relevant results.*

Key content: *Described are test item (characterisation, solubility and handling), test system and test item interaction and biokinetics during development to ensure test item compatibility and correct and reliable exposure.*

Guidance for improved practice: *More detailed information on solubility determination methods; the limitations of test items for which the method is suitable will allow the reader to choose the most suitable approach for his/her particular needs.*

Recommendations *Identify suitable reference and control items to avoid interference of the test, reference and/or control items with the in vitro method*

This chapter describes the characterisation and preparation of the test item and of relevant reference and control items for the *in vitro* method. Furthermore, details will be given as to how test item stability must be monitored, how the *in vitro* method environment can affect the test item and how the biokinetics of the test item in the *in vitro* method should thus be assessed. These aspects are important to ensure reproducibility among laboratories and certainty that the outcome of the *in vitro* method is indeed related to the test item. Additionally, it is described which reference and control items, such as negative and positive controls, should be applied in general to verify correct function of the *in vitro* method controls (OECD, 2004[1]).

A distinction is made between aspects that are important in the development phase of an *in vitro* method, and those that matter when the *in vitro* method is used routinely for regulatory purposes. During the phase of test development, chemicals or products with well-known characteristics should be tested to assess the relevance of the method and to obtain results which will be used to set the acceptance criteria (Section 6.1). In the development phase, it is important to determine:

1. Which items are suitable as reference and control items?
2. The applicability domain of the method, i.e. if the *in vitro* method can be used for liquids, solids, certain powders, mixtures (e.g. agrochemical formulations), multi-constituent substances, certain preparations, suspensions, nanoparticles, emulsions, etc. (OECD, 2002[2]).
3. Selection of the appropriate labelling method (if applicable) and the relative benefits/disadvantages (e.g., use of radioactive versus fluorescent labelling) and the potential of the test item to interact with the labelling method (e.g., phytoestrogens can interact with the luciferase and boost the signal, much beyond that for the endpoint in question).
4. The process of preparation or formulation of the test item, before applied to the *in vitro* method.
5. The concentration of solvent(s) that can be used without interfering with the *in vitro* method.
6. Limitations and uncertainties in the method.

Prior to routine use laboratory proficiency to perform the method will need to be shown (Section 8.4).

6.1. Reference and control items

The purpose of the reference item(s) is to grade the response of the test system to the test item, while the purpose of the control item(s) is to control the proper performance of the test system (OECD, 2004[1]). Since the purpose of control items may be considered as analogous to the purpose of a reference item, the definition of reference item may be regarded as covering the terms ‘positive, negative, and/or vehicle control items’. In this way it has been made clear that the definition of the reference item does not only include the use of an item used for the "absolute grading" of the response, but also for its use in "relative grading", i.e., the responsiveness of the test system (Seiler, 2005[3]).

Reference item(s) can be one or more item(s) where a specific readout and well-known response is expected (OECD, 2004[1]). The reference item(s) is used to provide a basis for comparison with the test item or to validate the response of the test system to the test item

i.e., provide a known measurable or observable response. Reference item(s) should be relevant to the endpoint being measured, have a well-defined chemical structure and purity (applicable only to chemical based reference items), should be non-hazardous (where possible) and should be available from commercial sources without prohibitive costs. Justification for the selection of the reference item(s) should be documented, preferably in the *in vitro* method SOP(s).

Reference item(s) should be tested for batch-to-batch variability and be appropriately characterised (e.g., purity, stability) and identified (e.g., Chemical Abstracts Service (CAS) number) (OECD, 1998[4]). Records of identity (CAS number, batch number, purity, chemical structure, molecular weight, etc.), receipt, storage, preparation and use should be available to allow for a full reconstruction of the history and use of all reference and control items.

Solubility, stability, and purity need to be established for each reference item used, and acceptance criteria based on historical data need to be developed. The continuous monitoring of the reference item(s), e.g., in the format of a control chart (Section 2.3), is important to prove that the *in vitro* method continues to perform within the set limits, and is consistent over time. It is recommended to use authentic standards or control the concentration of the reference item stock solution (preferably by a certified laboratory).

Control items are used to control the proper performance of the test system (OECD, 2004[1]) and therefore the validity of the executed experiments. The extent to which control items may need to be analytically characterised may differ from the requirements of reference items used for absolute grading (OECD, 2004[1]).

A negative control is an item is an item for which the test system should not give a response, while a vehicle (or solvent) control assures a response does not originate from the applied solvent. A positive control may also be used as a reference item for absolute grading of the response of the test item.

Selection of the positive control should begin as early in the *in vitro* method development process as is practical (Hartung et al., 2004[5]) as it can help identify dependent variables that impact the method consistency. Therefore, it is important that the positive control is run concurrently with the test item(s) each time the *in vitro* method is performed (Ulrey et al., 2015[6]). The ideal positive control item is one that has a consistent and predictive effect on the *in vitro* test system. As such, it needs to induce a known change in the endpoint measured and fall within the dynamic (quantifiable) range of the test, so that increased and decreased magnitudes of response can be measured.

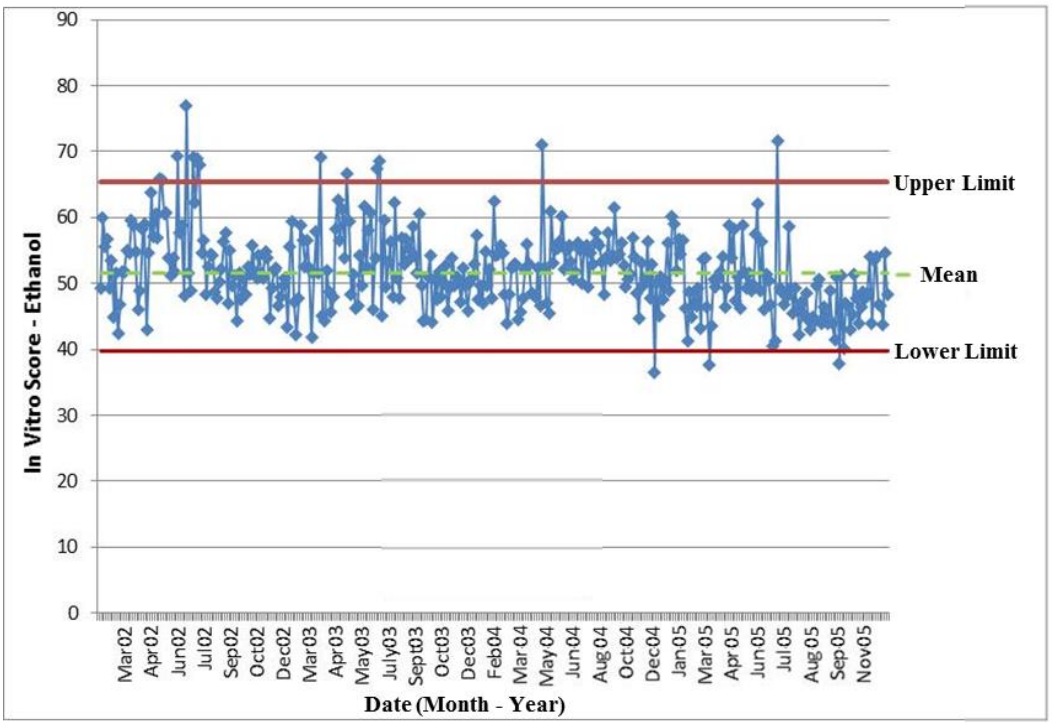
Monitoring and recording performance against negative and positive control items may constitute sufficient proof for the responsiveness of a given test system (OECD, 2004[1]). Non-response of the test system to the negative control and positive response to the positive control, within the acceptance criteria, show that the test system is "reactive" and behaves as expected. Responses outside of the expected range can be indicative of non-normal behaviour of the test system, e.g. due to a change in components of the test system, and could be a reason for a more thorough investigation of the test system responses. These acceptance criteria are usually based on the historical control values recorded in the test facility, which should furthermore be comparable to published reference values when available. Guidance on how to compile and use historical data can be found in literature, e.g., Hayashi (Hayashi et al., 2011[7]) describes the compilation and use of historical data specifically for genotoxicity data, but this approach can also be

applied in a broader context. A more general approach is described by Yoshimura (Yoshimura and Matsumoto, 1994[8]).

Sometimes it may be possible to select a single positive control to address all endpoints or exposure conditions. In genetic toxicity *in vitro* methods, such as the bacterial reverse mutation (Ames) assay (OECD, 1997[9]), two positive controls are used for each bacterial strain to address direct mutagenic activity and metabolic activation of a promutagen with rat liver S9 (Zeiger et al., 1988[10]). In the Bovine Cornea Opacity Permeability Test (BCOP), one positive control is used for the liquids exposure testing method and another for the solids exposure testing method (OECD, 2017[11]) (OECD, 2009).

Three historical examples are given below to exemplify the importance of including and monitoring a positive control concurrent with the *in vitro* method. The first example (Figure 6.1) is for the BCOP. It shows a quality control chart (Section 2.3) for the BCOP using ethanol as a positive control for each test performed over a period of two and a half years. The acceptable upper limit is between 60 and 70. As can be seen from the graph, there is a cluster of failed runs with values that are higher than expected, i.e. above the upper limit (in mid-2002). The basis of the failures was not immediately clear since the bovine eyes looked quite normal upon arrival in the laboratory. However, the pattern was persistent and the cause was eventually traced to improper handling of the eyes in the abattoir. Without the concurrent positive control data, it would not have been possible to identify the problem and prevent inappropriate data from being reported. Isolated tissues or tissue constructs as test systems can be difficult to properly evaluate visually and so the functional test provided by the concurrent controls is often the only way to measure their integrity.

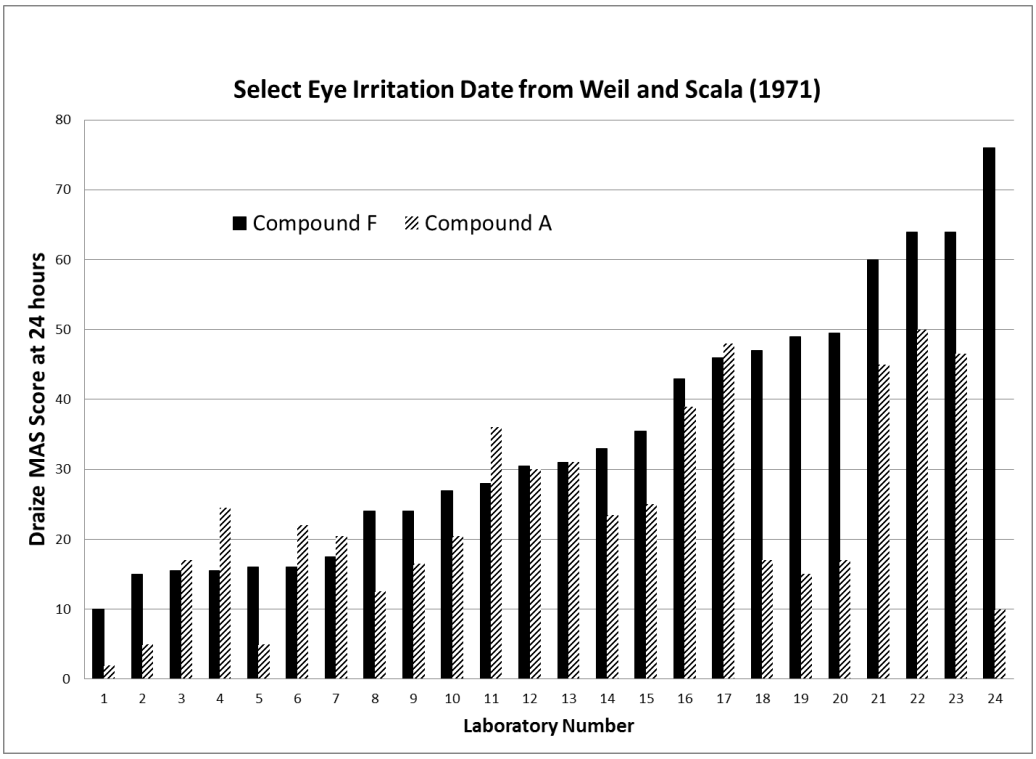
Figure 6.1. BCOP ethanol positive control QC chart March 2002 to November 2005



Source: courtesy of the Institute for *In Vitro* Sciences (IIVS), Maryland, USA.

The second example (Figure 6.2) shows the consequences of not including positive controls in each run as there was no means to compare intra- or inter-laboratory data. The data is obtained from a Draize eye irritation study, as published in "Study of intra- and inter-laboratory variability in the results of rabbit eye and skin irritation tests" (Weil and Scala, 1971[12]). It shows the 24-hour Draize Maximum Average Scores (MAS) scoring system for grading of ocular responses based on the rabbit eye model developed by Draize, which was the primary quantitative measurement of eye irritation potential in rabbits. Data is shown for 46% aqueous triethanolamine lauryl sulphate and 95% ethanol. The data are arrayed in order of increasing MAS values for 46% aqueous triethanolamine lauryl sulphate. The corresponding MAS for 95% ethanol is paired with the MAS for 46% aqueous triethanolamine lauryl sulphate from that laboratory, however no positive controls were used at that time.

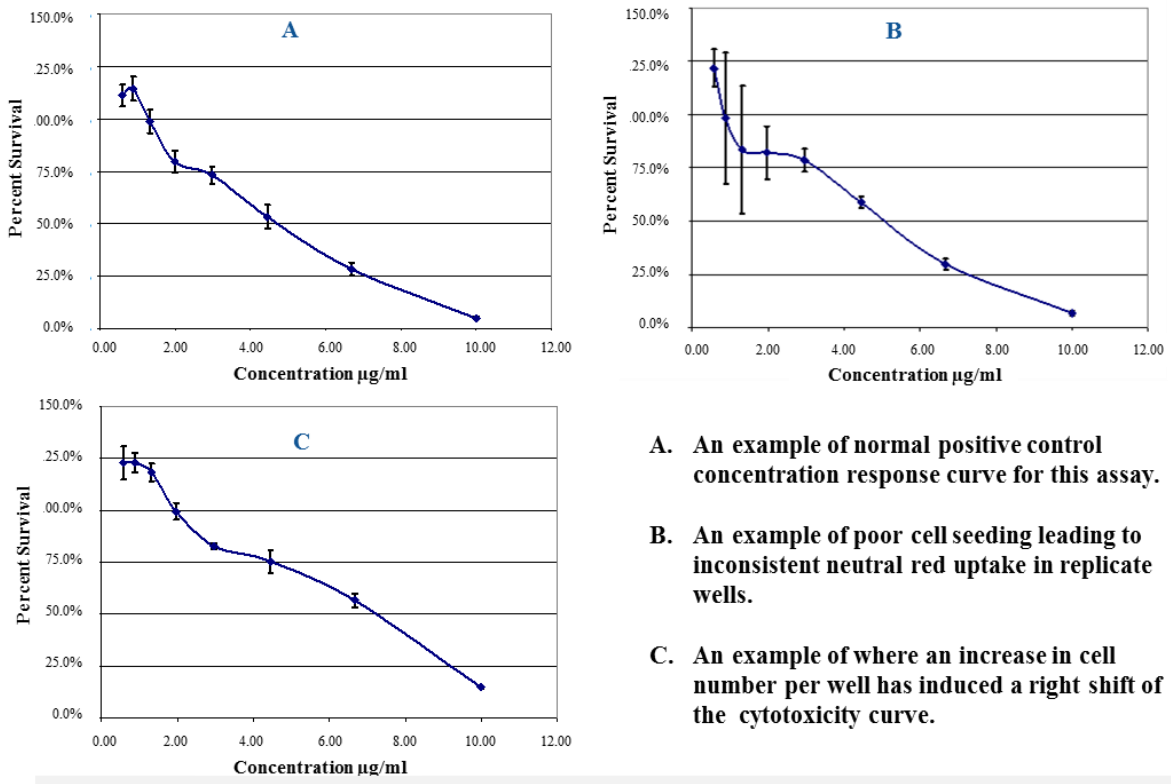
Figure 6.2. Historical data from the Draize Eye Irritation test on two chemicals at 24 hours after instillation where no positive control was included



Notes:
The data are arrayed in order of increasing Maximum Average Score (MAS) for compound F.
The corresponding MAS for compound A is paired with the MAS for compound F from that laboratory.

The third example of the relevance of a positive control is given in Figure 6.3. It shows three concentration response curves from a keratinocyte-based cytotoxicity assay (neutral red uptake endpoint) treated with sodium lauryl sulphate as a positive control. From the graphs it can be seen how *in vitro* method performance can be affected by a number of factors, e.g., due to errors in pipetting, which is apparent by looking at the (well characterised) dose-response curves of the positive control.

Figure 6.3. Neutral red uptake cytotoxicity assay using human keratinocytes



Three concentration response curves from a keratinocyte-based cytotoxicity assay (neutral red uptake endpoint) treated with sodium lauryl sulfate (also known as sodium dodecyl sulfate or SDS)

6.2. Applicability and limitations of the method

It is important to clearly describe the applicability domain of the *in vitro* method, as well as any limitations or exceptions. OECD Guidance Document 34 defines the applicability domain as *a description of the physicochemical or other properties of the substances for which a test method is applicable for use*. The applicability domain may also specify known limitations of the *in vitro* method such as restrictions on the classes of substances that can be accurately identified or measured by the *in vitro* method. In practice it is often easier to define the limitations of an *in vitro* method than to define the applicability domain based on a limited number of test items used during validation due to practical and economic reasons, i.e. only a limited number of test items can be assessed.

The applicability and limitations of the method need to be discerned in the development phase of the *in vitro* method. The *in vitro* method should include physical-chemical and other limitations using some or all of the properties listed below. The list is not exhaustive and may need to be further extended depending on the nature of the test item.

7. State: solid, liquid, gas, type of radiation and all in between-states such as aerosol, dust, or viscous liquid (OECD, 2012[13]). Depending on its state, the test item may require specific preparation steps before the test (Section 6.3) or a specific

administration mode in the method, such as dry dispersion with pressurised air, nebulisation of a liquid formulation, or spark generation.

8. Appearance: nominal size, morphology, size distribution, aggregation and agglomeration phenomena and surface characteristics (surface area, surface charge, surface chemistry) are essential characteristics to know the nature of a certain nanomaterial (OECD, 2012[14]).

9. Colour: some test items may interfere with the endpoint detection method if coloured test items are tested or coloured metabolites are generated.

10. Physicochemical characteristics (if available; some physicochemical property values may be experimental or predicted)

    - pH for test item in solution (OECD, 2013[15]) and the acid dissociation constant at logarithmic scale pKa (pKa indicates to what extent the test item may become ionised at the pH of the test system). Changes of pH can also affect the test item in other ways than its ionisation (OECD, 2004[16]).
    - Osmolality.
    - Volatility.
    - Solubility (Section 6.5).
    - Dissociation constants in water (OECD, 1981[17]): dissociation is the reversible splitting into two or more species which may be ionic. The dissociation governs the form of the test item in the test system, which in turn determines its behaviour and transport which may affect the adsorption of the substance to culture dishes or the penetration into cells or adsorption onto proteins in solution or resulting in aspecific aggregation behaviour.
    - Lipophilicity: determination of the partition coefficient i.e. K_{ow} (OECD, 2006[18]); (OECD, 1995[19]). Highly lipophilic substances tend to get "lost" in an *in vitro* system by adsorbing to the plasticware.
    - Homogeneity and conditions of stable homogeneity.
    - Fluorescence properties: interference due to autofluorescence or quenching.
    - Sensitivity to photolysis (OECD, 2008[20]).
    - Photoreactive potential (International Council on Harmonisation (ICH) S10 guidelines explains photoreactive potential of chemicals in relation to absorption of light with wave length of certain range).

11. Composition and purity (if available): chemical purity/contaminants, microbiological contaminants (including e.g., cell walls of decomposed microorganisms), biological purity (e.g., of cells lines or test microorganisms, or complex protein mixtures (vaccines)), composition of complexes (vegetal extracts, products of fermentation, etc.). In case of a mixed solution, the list of ingredients with percentages of each component can be relevant to describe the composition. For each component, information like molecular weight, chemical formula, CAS registration number, etc. is useful. Complex test items may require other information, e.g., Unknown or Variable composition, Complex reaction products or Biological materials (UVCBs) cannot be sufficiently identified by their chemical composition, because the number of constituents is relatively large

and/or the composition is, to a significant part, unknown and/or the variability of composition is relatively large or poorly predictable. The composition could then be defined by the manufacturing process description[1].

12. Conditions of stability (if available): the limits of temperature, pressure, and humidity to maintain stability of the test item (to be compared with the *in vitro* method conditions).
13. Microbiological status: aseptic conditions are always recommended to prevent unexpected effects by biological contaminants.
14. Sterility and expiry date: relevant for medical devices.

If the *in vitro* method is known to be amenable/not amenable to a variety of chemicals such as mixtures, UVCBs, multi-constituent substances, organometalics, inorganic substances, discrete organic substances and various chemical classes or organic substances (OECD, 2017[21]), these should also be described in the method. It is important to note that solubility is a highly important yet often neglected characteristic and is therefore described in more detail separately (Section 6.5).

6.3. Test item preparation

Test items may have to go through various steps of preparation, such as sterilisation, dissolution, dilution, extraction by wetting or centrifugation before being suitable for use in the *in vitro* method. The purpose of each step of the preparation has to be explained, and the critical limits of the step/procedure should be determined. The impact on the test item stability, homogeneity and integrity should be assessed, e.g. proper photo protection measures should be taken if it is relevant to the properties of the test item.

For more complex test items there are existing guidelines to aid this process: e.g., ISO 10993-12 gives extraction conditions needed to obtain a representative extract of medical devices depending on their composition, and the OECD series on the Safety of Manufactured Nanomaterials, n°36 (OECD, 2012[14]), gives advice on how to prepare and characterise a nanomaterial dispersion, while ISO/TS 19337:2016 describes characteristics of working suspensions of nano-objects for *in vitro* methods used to evaluate inherent nano-object toxicity.

The highest concentration of test item that should be tested may differ per *in vitro* method and needs to be defined in the study protocol/plan. Factors to be taken into consideration when deciding the highest concentration include the solubility and stability of the test item, its cytotoxicity, changes to the culture environment due to an increase or decrease in pH due to the test item, but may also be more specific, relating to the endpoint readout. The highest concentration may also be based on *in vivo* data if they exist (e.g., lowest dose at maximum effect *in vivo)*. The lowest concentration to be tested quite often will depend on the limits of quantification (LOQ) of the associated measuring instrumentation (Section 8.3.1), however this will ultimately depend on the concentration range of the test item to be tested.

6.4. Concentration range

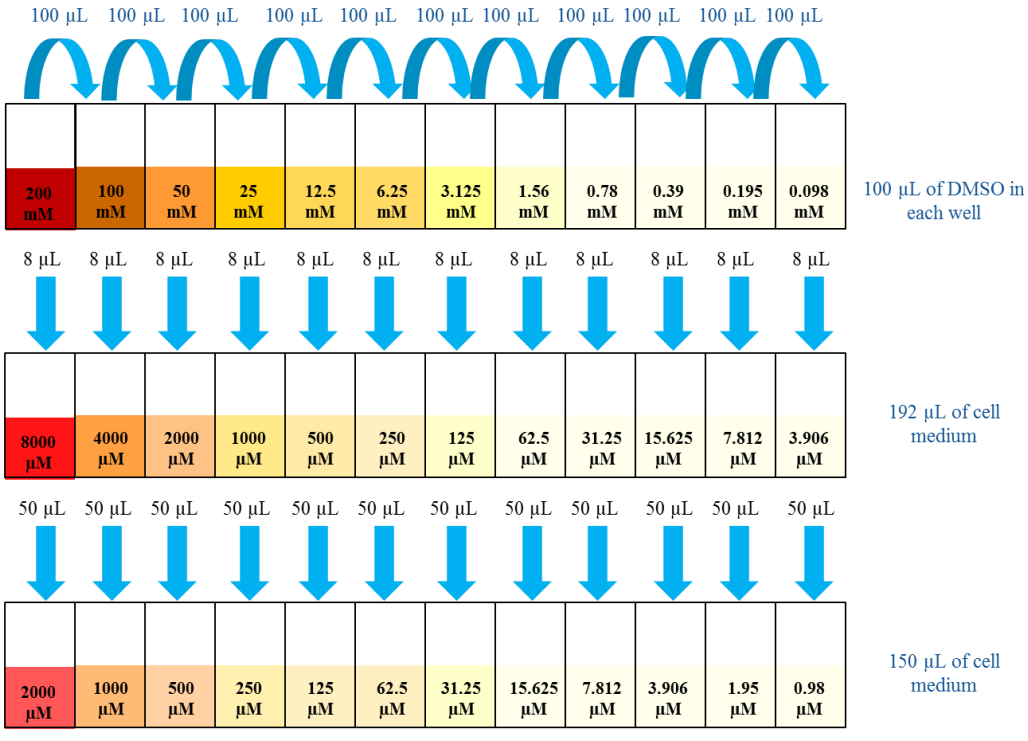
In many OECD TGs a preliminary test is carried out to determine the appropriate concentration (dose finding) range of the test item to be tested, and to ascertain whether the test item may have any solubility and cytotoxicity issues (e.g., TG 455, TG 442E).

The first run often tests the test item using log-serial dilutions starting at the maximum acceptable concentration (e.g., 1 mM, 100 µM, 10 µM, maximum solubility, etc.) to find a concentration-response curve. Further runs, using smaller serial dilutions (e.g., 1:2, 1:3) are used to focus in on the concentration-response curve, usually using six to eight concentrations (e.g., TG 442E). When a solvent is used, the maximum concentration which does not influence the test should be confirmed experimentally (e.g., DMSO 0.1% [v/v] for TG 455; organic solvents 1% [v/v] for TG 490).

To prepare a stock solution and the subsequent working solutions, serial or direct dilution methods can be used (Comley, 2007[22]). The serial dilution process increases the error in precision at each successive step; however direct dilution methods, which normally involve only one step, require specialised equipment.

An indicative procedure on how to prepare a series of working solutions from a stock solution of 200 mM is illustrated in Figure 6.4. It is recommended to use new tips for each step so as to eliminate any chance of carryover, specifically when pipetting from highest to lowest concentrations. However, multiple contacts also open the technique to problems caused by leachates[2]. After a series of diluted samples are generated (working solutions), a small, pre-determined volume is transferred to the *in vitro* method plate, increasing the potential for carryover. In this example, the final DMSO concentration in the *in vitro* method is 1.0 %. Note that 10 µL of sample is used to generate the concentration gradient.

Figure 6.4. Serial dilutions scheme for KeratinoSens™ test method



Source: (OECD, 2018[23])

Direct dilution follows a simpler process, where controlled volumes of the same concentrated stock solution are transferred directly to individual wells to achieve the

desired end concentration. Essential to direct dilution is the ability to accurately transfer extremely small volumes of stock solution, which is generally not possible with pipets. For this purpose, acoustic liquid handlers are frequently used in industry to generate concentration ranges via direct dilution. Because the amount of stock solution is so small, the sample can be maintained in pure DMSO to reduce the chance of sample precipitation during dilution. Compounding error is eliminated since samples are not serially diluted. In general, significantly less quantity is used to generate the final concentration ranges. Direct dilution also generates less liquid and solid waste, reducing expenses.

6.5. Solubility

Solubility defines how much of a substance (e.g., test, reference and/or control items), which for pure substances refers to its molecular and ionised forms, can be maximally dissolved in the solvent to be used for the *in vitro* method, while the rate of dissolving is called dissolution. Hence, the solubility value is a thermodynamic property while the dissolution rate is a kinetic one. In other words, time has no effect on the solubility value, but it is important in dissolution related items (Jouyban and A. Fakhree, 2012[24]).

Solubility is relevant because any precipitation would effectively lower the true concentration by an unknown amount to less than the nominal concentration prepared, and so influencing the significance and reproducibility of results. Precipitates may also affect read-outs of the *in vitro* method, e.g., increasing the optical density in absorbance measurements. For the majority of *in vitro* methods it is important to ensure that reference, control (where applicable) and test items are completely dissolved, however, there are some exceptions which do not require full solubility e.g., OECD TG 442D (OECD, 2018[23]), TG 442E (OECD, 2018[25]) and TG 487. In these cases, assuming other factors such as cytotoxicity, pH, osmolality, homogeneity, etc., are not predominant, test concentrations above the solubility limit may be applicable, with turbidity or slight precipitation present.

Solubility depends on the physical-chemical properties of the substance and on the type of solvent to be used. Furthermore, solubility is also affected by the composition of the substance (e.g., presence of impurities) and by the experimental conditions (temperature, incubation time, possible adsorption to the test vessel or to medium constituents such as plasma protein or albumin). Some general factors that may affect solubility are described in Table 6.1.

Table 6.1. Common factors affecting solubility

Factor	Affect
Temperature	In most cases solubility increases with temperature, with the exception of gases.
Polarity	In most cases, similar polarities of the solute and solvent increases solubility, e.g., polar solutes do not dissolve in non-polar solvents
Molecular size	As a general rule, excluding other factors, larger particles are generally less soluble
Agitation	Increases the speed of dissolving, i.e. dissolution
Ultrasonification	Increases the speed of dissolving, i.e. dissolution
pH	May affect the solubility of the solute
Pressure	Only affects the solubility of gases

Chemical test, reference and control items are generally dissolved in a solvent (e.g., DMSO, ethanol, purified water) to create a stock solution at a predetermined target concentration (e.g., stock concentration of 50mg/mL or 100mM). When selecting the

reference and control items, in the method development phase, their solubility and stability should be assured. Once the method is in use, test items that are not soluble in the desired solvent and/or concentration range should be considered incompatible with the *in vitro* method. The tolerable solvent concentrations will depend on the solvent and the test system used. As a general rule, the final solvent concentrations should be as low as possible to avoid any potential interference with the *in vitro* method, and may be as low as 0.1%. The final solvent concentration should be the same both for test item(s) and for the reference and control (where applicable) item(s) preparations.

In vivo methods are also frequently conducted by suspending insoluble test items in solvents, particularly for oral exposures. Experience has shown that insoluble particles dissolve over the course of time and digestion, and the same principle may also apply to *in vitro* exposure. Furthermore, for many *in vitro* methods, the measured response continues to increase with increasing test item concentration, also into the range at which precipitation occurs. This is particularly true of "well behaved suspensions" i.e., very fine particles which remain suspended rather than settling on to the cells and/or interfere with the *in vitro* method. Where applicable, an evaluation of this phenomenon should also be taken into consideration when assessing solubility. Possible interference of media components (e.g., serum, proteins, plastic, etc.) with the test item should also be considered as solubility of the test item may also be affected (Section Table 6.1).

Regarding nanomaterials, special issues on measuring solubility and dispersion characteristics may arise. Nanomaterials have special physicochemical and biological properties and they may agglomerate to form larger "particles" with properties different from the single nanoparticle. For these materials, the specific guidance documents[3] are best followed, which are continuously being developed (Scenhir , 2015[26]). Any toxicity testing using *in vitro* methods should pay special attention to the agglomeration/aggregation behaviour, and the insoluble/partially-soluble nature of nanomaterials (Scenhir , 2015[26]). Possibilities for dis-agglomeration and re-aggregation of nanomaterials should also be considered as some properties of nanomaterials may change due to interaction with the surrounding media.

6.5.1. Solubility determination

While computational methods provide solubility predictions for various solvent or matrices (Bergström et al., 2002[27]); (Persson et al., 2013[28]), they are not generally available for conditions specific to individual *in vitro* methods.

Visual inspection remains a simple and common approach to solubility determination, with HPLC/UV spectrophotometry and nephelometry generally applicable as instrumental analytical methods (Pan et al., 2001; Bevan & Lloyd, 2000; Hoelke et al., 2009) (Table 6.2).

Although subject to operator judgement, visual inspection can also be perceptive in assessment of solubility, enhanced by use of microscopy to detect particulate solid or immiscible liquid phase in suspension. Reliability can also be improved by centrifugation especially to determine precipitation in medium dilutions, where foaming may obscure visual observation.

OECD TG 105 (OECD, 1995[29]) can be used for the determination of aqueous solubility of pure substances which are stable in water and are not volatile, while OECD TG 116 (OECD, 1981[30]) can be used for fat solubility determination (fat solubility is the mass

fraction of substance which forms a homogeneous phase with a liquid fat (oil) without giving rise to chemical reactions).

- Nephelometry facilitates solubility determination, particularly suited to serial measurement (e.g., ranges of chemicals and/or concentrations) allowing systematic and precise evaluation of turbidity due to dispersed precipitation, independent of matrix composition. Furthermore, nephelometry can be used for preparations in biological media. However, the measurement is relative, requiring a definition of threshold turbidity for insolubility based on expedient practice with the detection limit dependent on instrument sensitivity. Moreover, nephelometry may not detect chemicals such as transparent immiscible liquids for which visual inspection, enhanced by experienced microscope observation, remains a reliable approach.
- UV spectrophotometry, LC/HPLC coupled with UV or MS methods provide a quantitative determination of the concentration with the use of standard curves. While these methods are valid for solutions prepared in solvent, they may not be valid for preparations in biological media, which contain many components that often interfere with analytical methods. Cell culture media cannot be injected into LC/HPLC columns and their multiple components will likely obscure the compound of interest through their inherent UV absorbance. This necessitates pre-purification and extraction steps to remove these components prior to the LC/HPLC step.

Table 6.2. Comparison between solubility determination methods

Method	Limitations	Specificity	Cut off	Rapidity
Nephelometry (Light scatter)	Sticky precipitates Impurities	Low	No	High
UV/VIS 1 (Absorbance)	Compound must have chromophore Sticky precipitates Impurities	Low	<500 nm	High
UV/VIS 1* (Filtration + Absorbance)	Compound must have chromophore Sticky precipitates Impurities Loss due to filter absorption	Medium	<250 nm	Medium
HPLC-UV*^	Sticky precipitates	High	No	Low
LC-MS*^	Sticky precipitates	High	No	Low
* Requires Calibration^ High Cost				

Solubility in both stock and working solutions (Section 6.7) should be determined. Regarding the sample preparation procedure, the following issues are key to producing reproducible results:

- Optimal time for dissolution in solvent: Does the test item dissolve immediately in the solvent or does it require additional treatment such as longer time frame, vortexing, sonication and/or heating.
- Solubility in media (Section 6.7): As the final dilution step usually involves transfer into the medium containing the test system, solubility should also be controlled upon transfer and include the incubation step, so as to mimic the *in vitro* method conditions, e.g., at the desired temperature and CO_2 levels and over the time period as described in the *in vitro* method.

- Visual inspection sample volume: Solutions for visual inspection should be prepared in a clear vial (e.g., glass) with a minimum volume of 0.5 ml, as smaller volumes are more difficult to assess with the naked eye.

In vitro method media typically have a rather high ionic strength and an inherently complex composition, which makes it difficult to predict the test item solubility upon dilution in the medium. It is therefore necessary to determine the solubility of the final concentration of the test item in the *in vitro* method medium under *in vitro* method conditions. In the case of inorganic substances, the anion and cation part of the test item may precipitate with other cations and anions present in the culture medium if the solubility limit of these newly combined salts is exceeded. It is therefore recommended to visually monitor the test item for precipitation as anions and cations present in the medium can form low-solubility salts with the test item.

6.6. Stability

The stability of test items under storage and in test conditions should be verified and expiry dates allocated as appropriate (OECD, 1998[4]). If a test item is administered in a solvent other than the standard buffer or cell culture medium, ideally, the homogeneity, concentration and stability of the test item in that solvent should be determined (OECD, 1998[4]). However, most *in vitro* methods procedures call for multiple runs and hence provide some measure for reproducibility of the results related to homogeneity and test item concentration.

The stability of the test item should be monitored throughout the exposure period as the concentration of the test item to which the test system is being exposed may vary with time. There are examples in literature available that describe compounds which have been hydrolysed in aqueous solutions (Crean et al., 2015[31]; Pomponio et al., 2015[32]).

6.7. Solvents

The compatibility of the solvent with the test system must be assessed, so as to select the appropriate solvent at an acceptable final concentration in the *in vitro* method medium. Strong toxic solvents with properties in terms of corrosivity, mutagenicity, carcinogenicity, genotoxicity or teratogenicity, which have the potential risk to induce adverse effects, should be avoided and only a compatible scale of solvents for stock solutions preparation should be considered. Another consideration to take into account is the possible masking of the *in vitro* response due to interference of the solvent with the test system (Coecke et al., 2016[33]).

The solubility limit of a test item will depend on the solvent of choice (among other things). DMSO is a commonly used solvent as it dissolves both polar and non-polar compounds and is miscible with a wide range of organic solvents and water. However, DMSO is highly hygroscopic and rapidly absorbs water in air which may result in a change of test item concentration.

As a common practice, organic solvents are generally used to prepare the stock concentration even if the test item can also be dissolved in purified water. One of the reasons is that organic solvents prevent or minimise the growth of microorganisms which can then impact the test item stability over time. OECD TG 455 recommends that the test item(s) should be dissolved in a solvent that solubilises that test item and is miscible with the cell medium, e.g., water, ethanol (95% to 100% purity) and DMSO. Other solvents

and concentrations may be used for specific methods as directed by test guidelines and/or study documents (e.g., Study Plan or SOP(s)). In most cases the test item should have a relatively high solubility in the solvent of choice, at a minimum producing a workable suspension, and the solvent should not interfere with the test item (e.g., inactivate the compound) or test system. For example, DMSO can reduce the effects of platinum complexes (Hall et al., 2014[34]). In addition, the solvent should not affect cell health or the phenotype of the cells used in the *in vitro* method when diluted in media and its concentration should be kept as low as possible. Common solvent concentrations for DMSO and ethanol are ≤1% (defined in individual methods or TGs), though 100% acetone is used for skin genotoxicity *in vitro* methods (Meza-Zepeda et al., 2008[35]). Toxicity of the solvent to the test system should be assessed by comparing the untreated control response with the solvent control response.

Insoluble test items may require more specific solvents, e.g., in OECD TG 442E (IL-8 Luc *in vitro* method) a serum-free medium is proposed as a solvent while mineral oil is recommended in OECD TG 491.

6.8. Air-liquid-interface exposure

For certain types of organs, such as skin, eyes and lungs, the use of cultures at the air-liquid interface reflects human conditions more closely (Ahmad et al., 2014[36]); (Jean et al., 2010[37]); (Li, Ootani and Kuo, 2016[38]). Some test systems, using non-transformed human keratinocytes to reconstruct the skin epithelium based on the air-liquid interface cell culture technique, are used for *in vitro* skin irritation testing. With this technique multiple layers of viable epithelial cells (basal layer, *stratum spinosum, stratum granulosum*) are recreated under a functional *stratum corneum* (OECD, 2015[39]). *In vitro* exposure methods for addressing inhalation toxicity on cultured human lung cells conventionally rely on prior suspension of particles in a liquid medium. Such exposure systems have limitations and may modify the particle composition. Other techniques such as electrostatic precipitation can be used (de Bruijne et al., 2009[40]). It is important when using air-liquid culture systems for inhalation toxicity applications that standards/criteria used for the generation of aerosols of the test item(s) should be included and must be appropriate for the specific cell culture conditions (Lenz et al., 2014[41]).

6.9. Biokinetics/dose extrapolation/interference with media components

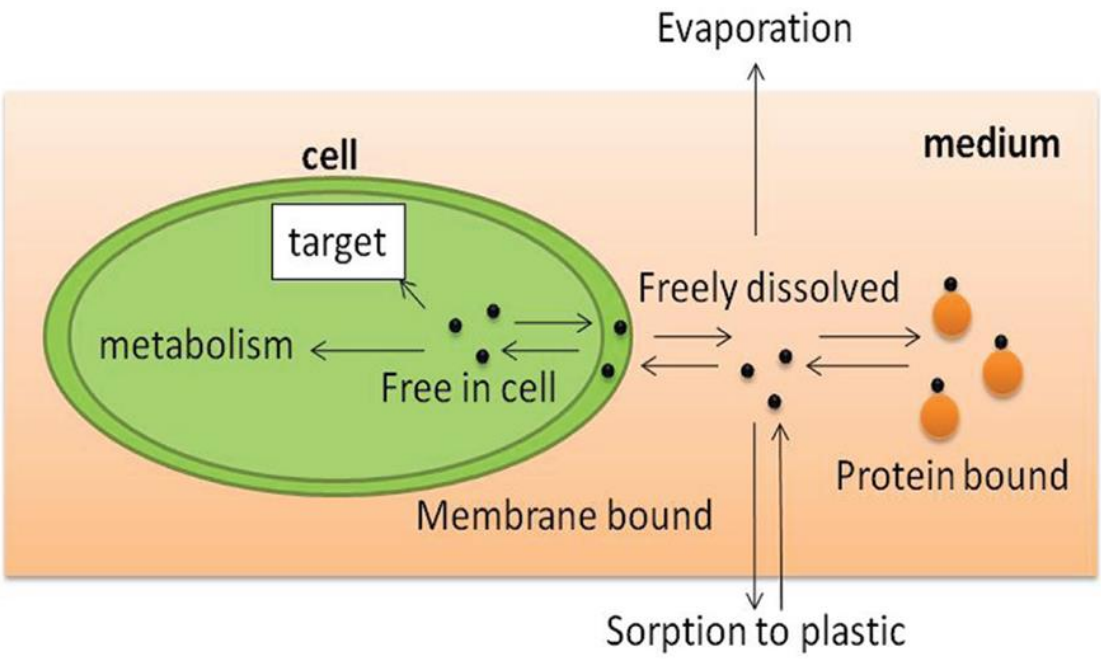
Just like the biokinetics *in vivo* are about what the body of the organism does to the test item, the biokinetics *in vitro* concerns what the *in vitro* test environment does to the test item. A central issue in biokinetics is that generally only the freely dissolved molecules of a chemical can pass the membrane barriers and reach a target inside a cell. Thus, in an *in vitro* system, the freely dissolved concentration of the test item in the medium or in the cell (being as close to the target as possible) is the central parameter. Some processes (Groothuis et al., 2015[42]; Heringa et al., 2006[43]); depicted schematically in , result in a freely dissolved concentration that is not the same as the nominal concentration (i.e., the added concentration).

A description of processes which affect xenobiotic *in vitro* bioavailability, change the identity of the test item and affect test item stability are described in Annex G: Biokinetics and xenobiotic bioavailability. In short, the free concentration can be decreased by evaporation of the test item, by adsorption of the test item to the plastic well of the culture plate, by binding to serum proteins in the culture medium, by absorption in

the lipid-rich cell membrane, by hydrolysis, by photolysis and by enzymatic metabolism in the cells. Regarding the binding of serum proteins including specific biological factors the following points should be taken into consideration when using serum:

1) if the test item is known to bind to protein, its effect might not be seen unless a very high concentration of test item is used (Section 6.9 on biokinetic parameters);
2) if the test item antagonises an endogenous circulating hormone or factor, the serum might contain such hormone or factor and may thus affect the *in vitro* method results.

Figure 6.5. Schematic representation of some processes that can cause the final target concentration to be different than the nominal concentration in an *in vitro* test



Source: Kramer *et al.,* 2012

Considerable variation between test facilities may be obtained where results (e.g., EC_{50} values) are based on the added, or nominal, concentrations, and as such may be unfit for extrapolation to *in vivo* (Kramer et al., 2015[44]), e.g., if there is considerable evaporation, the EC_{50} *in vitro* will appear to be much higher than it will be in the same tissue *in vivo*. Thus, in order to obtain pure EC_{50} values, that relate target concentrations to responses, these target concentrations should be measured. An EC_{50} value (based on a nominal concentration applied) determined for a specific test system may be used for ranking of test items on a specific effect. Due to low complexity of *in vitro* systems compared to *in vivo* a direct extrapolation of *in vitro* measurements to *in vivo* is often not feasible. For extrapolation of data also reliable *in vivo* data on target organ concentrations are needed for validation. Thus, determination of free concentrations in the cell or in the culture medium may be helpful but they are not mandatory for all test systems. Annex G: Biokinetics and xenobiotic bioavailability describes when and how such measurements can be performed and how *in vitro* results can subsequently be extrapolated to the *in vivo* situation.

6.10. Interference with the test system

A two-way interaction has to be assumed between the test item and the test system. In one respect, the test system can affect the test item (in analogy to biokinetics in *in vivo* models; detailed in Section 6.9). In another respect, the test item can affect the test system in specific ways (alteration of a readout) in accordance with the design and intended application of the test system (Chapter 5 and Chapter 8) or in unintended ways, by interfering with the overall performance of the biological model on which the test system is based, or by disturbing the readout of the *in vitro* method endpoint. There are endless possibilities for artefacts to be created in this way. As not all of these can be controlled for, experienced operators and personnel interpreting the *in vitro* method data are required to detect potential problems. Problem detection is also facilitated by regular inclusion of consistency controls and plausibility considerations (e.g., do compounds with similar structure or similar mode of action behave similarly? Can effects be reversed? Does another test system for the same biological process give similar results? Are findings consistent with biological expectations concerning concentration and timing of effect?).

Test items can disturb the test system, especially if it is based on living cells, as they are highly responsive to changes in their environment. The most frequent and serious disturbance is general cytotoxicity often leading to cell death.

6.10.1. Cytotoxicity

While the strict definition of cytotoxicity refers to cell death, a wider interpretation also includes adverse effects on cells that alter their functionality but do not lead to cell death (within the observation period). For instance, protein synthesis may be impaired, or mitochondrial function altered. Cytostasis, where dividing cells do not die but cease dividing, is another example of delayed cell death which can impact the endpoint measures. This can affect the specific endpoint of a test system (e.g., expression of reporter enzyme or speed of proliferation), without being relevant for the intended *in vitro* method objective.

Cytotoxicity is a useful and widely used marker for setting concentration levels for the test item, as well as providing some information of interaction between the cells in culture and the test item. However, if the *in vitro* method endpoint is not cytotoxicity, then cytotoxicity triggered by a test item is a serious confounder and needs to be controlled for. Indeed, changing cell numbers *in vitro* is known to affect observed effect concentrations (Gülden, Kähler and Seibert, 2015[45]); (Gülden, Mörchel and Seibert, 2001[46]) and this is particularly critical in repeated treatments (Kramer et al., 2015[44]). Therefore, it is important to understand the kinetics and the mechanisms of cytotoxicity as a simple type of calibrator to investigate the effects of incubation time, dose of test item, and plating density in cell-based cytotoxicity assays (Riss and Moravec, 2004[47]).

Measurement of cytotoxicity should be done using the same conditions as used for the specific *in vitro* method endpoint (i.e., in identical samples, ideally during the same run, or even better on the same plate), so as to obtain reliable and relevant cytotoxicity data. Alternative approaches use measurements in parallel cultures. Viability controls in related, but not identical, culture conditions (different plate format, different cell preparation, etc.) should be avoided. The choice of method used for cytotoxicity determination (Annex I: List of viability testing methods (non-inclusive) of cell cultures) but also the interpretation and reporting of the results needs careful consideration. It is important that the type of cytotoxicity method chosen is appropriate for use with the *in*

vitro method, specifically with regards to the timing of the *in vitro* method endpoint (e.g., if the *in vitro* method endpoint is performed after 4 hours, the cytotoxicity method should also be relevant at the 4 hours' time-point). Where the timing of the cytotoxicity method endpoint measurement does not coincide with the time of the *in vitro* method endpoint determination, the cytotoxicity measurement may be modified to coincide with the method endpoint. In this case it is important to verify that the cytotoxicity method performs as expected with the inclusion of positive and negative controls. There are no established rules on how to deal with this (relatively frequent) situation. One solution is to follow up on results from alternative methods for the same endpoint, or by using the same method with a changed incubation scheme (e.g., prolonged incubation). This is particularly important if data are used for risk assessment and far-reaching regulatory decisions.

A single endpoint is usually not sufficient to be fully conclusive. A combination of cell counting and a population measurement (e.g., resazurin reduction), or a combination of a viability measurement (e.g., calcein staining, dye exclusion, neutral red uptake) and a cell death measurement (e.g., propidium iodide uptake, Lactate DeHydrogenase (LDH)-release, Annexin V staining) provides a greater level of certainty. Importantly, controls for the viability measurement should be included and need to be considered for normalisation of viability data. For cells transfected with a plasmid encoding for a mutated androgen receptor that is constitutively expressed for cytotoxicity measurements, Alamar blue has proven an ideal cytotoxicity test as it can be directly compared with the response from the wild type receptor (Vinggaard, Bonefeld Joergensen and Larsen, 1999[48]).

6.10.2. Disturbed differentiation state

A special case of artefacts caused by test items is the change of biological properties of the test system without overt cytotoxicity. The most common example is an altered differentiation of cells or an altered composition of cell sub-populations. For instance, a test item might disturb cell differentiation state (Fritsche et al., 2005[49]) in a migration assay, and this alteration might lead to altered migration (Miettinen et al., 2000[50]). The item would be wrongly classified as modifying cell migration. Another example would be measurement of monocyte function (e.g., cytokine release) in a whole blood assay. If a test item leads to platelet degranulation, it might influence the overall endpoint of the *in vitro* method without affecting the monocyte response as such.

6.10.3. Altered communication/adhesion properties

Another special case of artefacts can be generated by interference of the test item with cell adhesion or communication. This is listed here separately, as it would not normally be detected by cytotoxicity assays, but it would strongly alter the behaviour of the test system (biological model) in the *in vitro* method. An example is binding of test item to molecules used for the coating of culture dishes. This would then alter readouts such as migration or neurite growth, without really affecting such processes within the cells (and without necessarily being relevant *in vivo*).

6.11. Interferences in the *in vitro* method

For pure test items, most of the unwanted interactions with the *in vitro* method are due to undesired interactions with either the test system or the *in vitro* method endpoint. The majority of interferences with the *in vitro* method endpoint are related to cytotoxicity

(immediate or delayed cell death or functional impairment), as covered in Section 6.10.1. The situation regarding test items containing impurities or non-inert additional substances in their formulation (Section 6.12) is more complex, and highlights the need to have clear specification for the test item. For instance, impurities (e.g., detergents or solvents) may alter skin or blood-brain-barrier (BBB) permeability (without being cytotoxic) and thus result in incorrect data on the test item of interest when measuring skin or BBB permeation capacity. In other cases, where the test item is a finished product, potential impurities and contaminations are considered part of the test item and their effect on the response is important and has to be evaluated.

Interference of the test item with the analytical endpoint means that the test item disturbs the normal measurement results (Thorne, Auld and Inglese, 2010[51]). This can be controlled by performing the *in vitro* method using adequate positive, negative, untreated or solvent controls. If the endpoints are of an analytical nature, the controls can also be spiked with the test item to verify that the test item does not in any way hinder the normal function of the test system or interfere with the readout.

Examples of read-out specific interference include, but are not limited to the following (Thorne, Auld and Inglese, 2010[51]).

Fluorescence/absorbance-based methods: disturbance by test items which are fluorescent or absorb light at the wavelength of measurement, or test items that quench fluorescence, or light scattering due to e.g., insolubility or bubble formation.

Luciferase based methods: non-specific activation or inhibition of the luciferase signal that can occur in a concentration-dependent manner.

Enzymatic assays: alteration of enzyme function, of co-factor, or of other limiting reagents by the test item; display of enzymatic activity (or chemical reactivity) by test item itself.

Resazurin or MTT reduction: strongly reducing agents directly reduce resazurin or 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) non-enzymatically. Compounds that trigger the release of superoxide can trigger reduction of resazurin by superoxide. This results in erroneous cytotoxicity data. Coloured compounds may interfere with the MTT measurement depending on the concentration and extinction coefficient.

Another relevant example of this kind of interference is provided by the interactions between reagents and nanomaterials in colorimetric assays for cytotoxicity (such as sulforhodamine B dye, or MTT used in the viability assays) (Scenhir , 2015[26]). Moreover, some nanomaterials may themselves disperse/absorb light and therefore interfere with the measurements in colorimetric assays. Some of these problems might be overcome by either adding appropriate controls or modifying existing protocols, e.g., removal of nanomaterials via centrifugation before reading the assay can reduce the variations in data generated for the same nanomaterials (Scenhir , 2015[26]).

6.12. Consideration of interferences not coming from the active ingredient

With test items that are not pure, interferences with the *in vitro* method may be caused by impurities or ingredients of the formulation. Particularly difficult cases arise when such additional chemicals are inactive alone, but synergize somehow with the effect of the test item.

This can also occur for the solvent of the test item. Frequently, a solvent concentration that does not affect the standard endpoint of an *in vitro* method as such (e.g., 0.1 or 1% DMSO) may still alter the effect of a test item on the test system (e.g., in the case of DMSO: through the antioxidant properties of DMSO; or through its effect on cell membranes; or through other activities including cell differentiation).

For test items consisting of a natural mixture (e.g., essential oils) or non-natural/artificial mixtures (e.g., agrochemical formulations), it should be considered to test the mixture as well as the known pure substances present, since the other ingredients of the mixture can change the overall effect of the test item. The individual kinetics of the ingredients must then be considered, although ingredients that are not absorbed *in vivo* will not have an effect on the test item systemic toxicity, they may however affect test item toxicity *in vitro*.

Notes

1. See: https://echa.europa.eu/documents/10162/22816103/10_sb_siduvcb_d1_lrws_20120203_en.pdf
2. See: https://en.wikipedia.org/wiki/Leachate
3. See: http://www.oecd.org/env/ehs/nanosafety/publications-series-safety-manufactured-nanomaterials.htm

References

Ahmad, S. et al. (2014), "In vitro Cell Culture Model for Toxic Inhaled Chemical Testing", *Journal of Visualized Experiments* 87, http://dx.doi.org/10.3791/51539. [36]

Bergström, C. et al. (2002), , *Pharmaceutical Research*, Vol. 19/2, pp. 182-188, http://dx.doi.org/10.1023/a:1014224900524. [27]

Coecke, S. et al. (2016), "Practical Aspects of Designing and Conducting Validation Studies Involving Multi-study Trials", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_5. [33]

Comley, J. (2007), "Serial vs direct dilution - Time to apply new thinking to IC50 determination and dose-response analysis?", *Drug Discovery World*, Vol. 8. [22]

Crean, D. et al. (2015), "Development of an in vitro renal epithelial disease state model for xenobiotic toxicity testing", *Toxicology in Vitro*, Vol. 30/1, pp. 128-137, http://dx.doi.org/10.1016/j.tiv.2014.11.015. [31]

de Bruijne, K. et al. (2009), "Design and Testing of Electrostatic AerosolIn VitroExposure System (EAVES): An Alternative Exposure System for Particles", *Inhalation Toxicology*, Vol. 21/2, pp. 91-101, http://dx.doi.org/10.1080/08958370802166035. [40]

Fritsche, E. et al. (2005), "Polychlorinated Biphenyls Disturb Differentiation of Normal Human Neural Progenitor Cells: Clue for Involvement of Thyroid Hormone Receptors", *Environmental Health Perspectives*, Vol. 113/7, pp. 871-876, http://dx.doi.org/10.1289/ehp.7793. [49]

Groothuis, F. et al. (2015), "Dose metric considerations in in vitro assays to improve quantitative in vitro–in vivo dose extrapolations", *Toxicology*, Vol. 332, pp. 30-40, http://dx.doi.org/10.1016/j.tox.2013.08.012. [42]

Gülden, M., D. Kähler and H. Seibert (2015), "Incipient cytotoxicity: A time-independent measure of cytotoxic potency in vitro", *Toxicology*, Vol. 335, pp. 35-45, http://dx.doi.org/10.1016/j.tox.2015.07.002. [45]

Gülden, M., S. Mörchel and H. Seibert (2001), "Factors influencing nominal effective concentrations of chemical compounds in vitro: cell concentration", *Toxicology in Vitro*, Vol. 15/3, pp. 233-243, http://dx.doi.org/10.1016/s0887-2333(01)00008-x. [46]

Hall, M. et al. (2014), "Say No to DMSO: Dimethylsulfoxide Inactivates Cisplatin, Carboplatin, and Other Platinum Complexes", *Cancer Research*, Vol. 74/14, pp. 3913-3922, http://dx.doi.org/10.1158/0008-5472.can-14-0247. [34]

Hartung, T. et al. (2004), *A modular approach to the ECVAM principles on test validity*, ATLA Alternatives to Laboratory Animals. [5]

Hayashi, M. et al. (2011), "Compilation and use of genetic toxicity historical control data", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 723/2, pp. 87-90, http://dx.doi.org/10.1016/j.mrgentox.2010.09.007. [7]

Heringa, M. et al. (2006), "Measurement of the free concentration of octylphenol in biological samples with negligible depletion-solid phase microextraction (nd-SPME): Analysis of matrix effects", *Journal of Chromatography B*, Vol. 834/1-2, pp. 35-41, http://dx.doi.org/10.1016/j.jchromb.2006.02.009. [43]

Jean, J. et al. (2010), "Effects of Serum-Free Culture at the Air–Liquid Interface in a Human Tissue-Engineered Skin Substitute", *Tissue Engineering Part A*, Vol. 17/7-8, pp. 877-888, http://dx.doi.org/10.1089/ten.tea.2010.0256. [37]

Jouyban, A. and M. A. Fakhree (2012), "Experimental and Computational Methods Pertaining to Drug Solubility", in *Toxicity and Drug Testing*, InTech, http://dx.doi.org/10.5772/19206. [24]

Kramer, N. et al. (2015), "Biokinetics in repeated-dosing in vitro drug toxicity studies", *Toxicology in Vitro*, Vol. 30/1, pp. 217-224, http://dx.doi.org/10.1016/j.tiv.2015.09.005. [44]

Lenz, A. et al. (2014), "Efficient Bioactive Delivery of Aerosolized Drugs to Human Pulmonary Epithelial Cells Cultured in Air–Liquid Interface Conditions", *American Journal of Respiratory Cell and Molecular Biology*, Vol. 51/4, pp. 526-535, http://dx.doi.org/10.1165/rcmb.2013-0479oc. [41]

Li, X., A. Ootani and C. Kuo (2016), "An Air–Liquid Interface Culture System for 3D Organoid Culture of Diverse Primary Gastrointestinal Tissues", in *Methods in Molecular Biology, Gastrointestinal Physiology and Diseases*, Springer New York, New York, NY, http://dx.doi.org/10.1007/978-1-4939-3603-8_4. [38]

Meza-Zepeda, L. et al. (2008), "High-resolution analysis of genetic stability of human adipose tissue stem cells cultured to senescence", *Journal of Cellular and Molecular Medicine*, Vol. 12/2, pp. 553-563, http://dx.doi.org/10.1111/j.1582-4934.2007.00146.x. [35]

Miettinen, P. et al. (2000), "Impaired migration and delayed differentiation of pancreatic islet cells in mice lacking EGF-receptors", *Development*, Vol. 127, pp. 2617-2627. [50]

OECD (2018), *Test No. 442D: In Vitro Skin Sensitisation: ARE-Nrf2 Luciferase Test Method*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264229822-en. [23]

OECD (2018), *Test No. 442E: In Vitro Skin Sensitisation: In Vitro Skin Sensitisation assays addressing the Key Event on activation of dendritic cells on the Adverse Outcome Pathway for Skin Sensitisation*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264264359-en. [25]

OECD (2017), *Guidance Document for Describing Non-Guideline In Vitro Test Methods*, OECD Series on Testing and Assessment, No. 211, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274730-en. [21]

OECD (2017), *Test No. 437: Bovine Corneal Opacity and Permeability Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264203846-en. [11]

OECD (2015), *Test No. 439: In Vitro Skin Irritation: Reconstructed Human Epidermis Test Method*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264242845-en. [39]

OECD (2013), *Test No. 122: Determination of pH, Acidity and Alkalinity*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264203686-en. [15]

OECD (2012), *Guidance on Sample Preparation and Dosimetry for the Safety Testing of Manufactured Nanomaterials*, OECD Publishing, Paris. [14]

OECD (2012), *Test No. 114: Viscosity of Liquids*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264185180-en. [13]

OECD (2008), *Test No. 316: Phototransformation of Chemicals in Water – Direct Photolysis*, OECD Guidelines for the Testing of Chemicals, Section 3, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264067585-en. [20]

OECD (2006), *Test No. 123: Partition Coefficient (1-Octanol/Water): Slow-Stirring Method*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264015845-en. [18]

OECD (2004), *Test No. 111: Hydrolysis as a Function of pH*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069701-en. [16]

OECD (2004), *The Application of the Principles of GLP to in vitro Studies*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 14, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084971-en. [1]

OECD (2002), *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures*, OECD Series on Testing and Assessment, No. 23, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078406-en. [2]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [4]

OECD (1997), *Test No. 471: Bacterial Reverse Mutation Test*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264071247-en. [9]

OECD (1995), *Test No. 105: Water Solubility*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069589-en. [29]

OECD (1995), *Test No. 107: Partition Coefficient (n-octanol/water): Shake Flask Method*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069626-en. [19]

OECD (1981), *Test No. 112: Dissociation Constants in Water*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069725-en. [17]

OECD (1981), *Test No. 116: Fat Solubility of Solid and Liquid Substances*, OECD Guidelines for the Testing of Chemicals, Section 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069800-en. [30]

Persson, L. et al. (2013), "Computational Prediction of Drug Solubility in Lipid Based Formulation Excipients", *Pharmaceutical Research*, Vol. 30/12, pp. 3225-3237, http://dx.doi.org/10.1007/s11095-013-1083-7. [28]

Pomponio, G. et al. (2015), "In vitro kinetics of amiodarone and its major metabolite in two human liver cell models after acute and repeated treatments", *Toxicology in Vitro*, Vol. 30/1, pp. 36-51, http://dx.doi.org/10.1016/j.tiv.2014.12.012. [32]

Riss, T. and R. Moravec (2004), "Use of Multiple Assay Endpoints to Investigate the Effects of Incubation Time, Dose of Toxin, and Plating Density in Cell-Based Cytotoxicity Assays", *ASSAY and Drug Development Technologies*, Vol. 2/1, pp. 51-62, http://dx.doi.org/10.1089/154065804322966315. [47]

Scenhir (2015), *Guidance on the Determination of Potential Health Effects of Nanomaterials Used in Medical Devices*, EU Publications. [26]

Seiler, J. (2005), *Good Laboratory Practice — the Why and the How*, Springer-Verlag, Berlin/Heidelberg, http://dx.doi.org/10.1007/3-540-28234-3. [3]

Thorne, N., D. Auld and J. Inglese (2010), "Apparent activity in high-throughput screening: origins of compound-dependent assay interference", *Current Opinion in Chemical Biology*, Vol. 14/3, pp. 315-324, http://dx.doi.org/10.1016/j.cbpa.2010.03.020. [51]

Ulrey, A. et al. (2015), "Applying Good Laboratory Practices (GLPs) to in vitro studies, one laboratory's perspective", *Institute for In Vitro Sciences*. [6]

Vinggaard, A., E. Bonefeld Joergensen and J. Larsen (1999), "Rapid and Sensitive Reporter Gene Assays for Detection of Antiandrogenic and Estrogenic Effects of Environmental Chemicals", *Toxicology and Applied Pharmacology*, Vol. 155/2, pp. 150-160, http://dx.doi.org/10.1006/taap.1998.8598. [48]

Weil, C. and R. Scala (1971), "Study of intra- and interlaboratory variability in the results of rabbit eye and skin irritation tests", *Toxicology and Applied Pharmacology*, Vol. 19/2, pp. 276-360, http://dx.doi.org/10.1016/0041-008x(71)90112-8. [12]

Yoshimura, I. and K. Matsumoto (1994), "Notes on the use of historical controls.", *Environmental Health Perspectives*, Vol. 102/suppl 1, pp. 19-23, http://dx.doi.org/10.1289/ehp.94102s119. [8]

Zeiger, E. et al. (1988), "Salmonella mutagenicity tests: IV. Results from the testing of 300 chemicals", *Environmental and Molecular Mutagenesis*, Vol. 11/S12, pp. 1-18, http://dx.doi.org/10.1002/em.2850110602. [10]

Chapter 7. Standard operating procedures

Key message: *Standard Operating Procedures (SOPs) and the accompanying forms, templates or worksheets should be written and prepared in a way that they will form the tools to simplify the work of the user when carrying out an in vitro method study.*

Key content: *The chapter elaborates on the correct documentation of in vitro methods for routine testing including requirements for clear and concise SOPs.*

Guidance for improved practice: *The evolution of a non-routine in vitro method to a routine in vitro method is described in a step-wise manner.*

Recommendations *to derive a set of clear, well-written in vitro method SOPs are given.*

According to the OECD Principles of Good Laboratory Practice (GLP) (OECD, 1998[1]), Standard Operating Procedures (SOPs) are defined as documented procedures which describe how to perform testing methods or activities normally not specified in detail in study plans or Test Guidelines (TGs). Formal SOPs facilitate consistency in the quality and integrity of a product or end-result, and are required by GLP. SOPs may include testing methods, instructions, worksheets, and laboratory operating procedures. SOPs are essential in a quality management system and must be formally authorised by management in a GLP test facility.

The aim of SOPs is to ensure that procedures are carried out in a consistent and reproducible way by qualified personnel. Therefore SOPs need to describe, in sufficient detail, clear work instructions for a trained user to minimise the risk for misinterpretation.

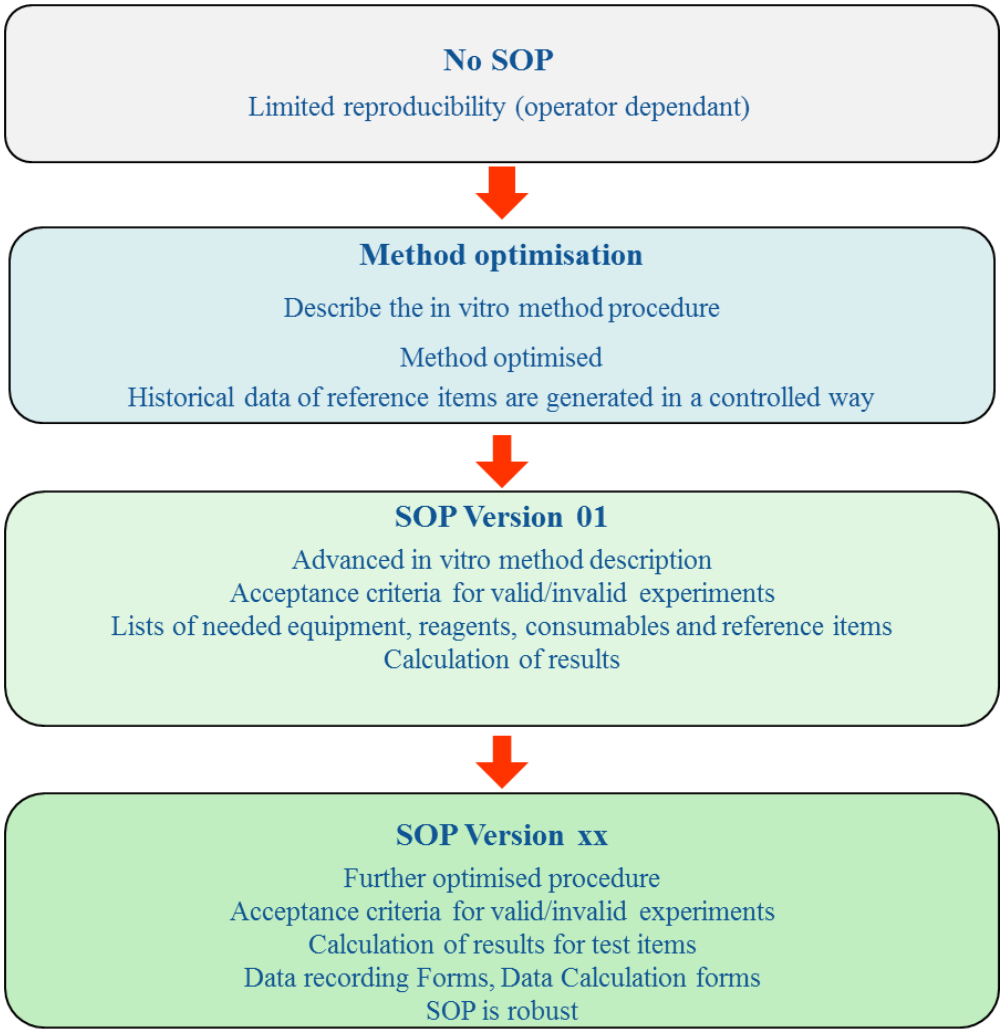
An *in vitro* method will be supported and documented with a number of different SOPs, forms, templates and worksheets. Besides the description of the main test procedure, SOPs for supporting procedures (e.g., the handling of cell cultures, waste handling, cleaning procedures, operating and calibration instructions for the equipment, record keeping, reporting, archival, quality assurance procedures, etc.) need also to be available and used. To avoid lengthy documents, the instructions are preferably divided into a series of SOPs. The SOPs must be readily available to personnel in each working area.

7.1. *In vitro* method standard operating procedures development

The development of an in vitro method for regulatory testing purposes is a difficult and time-consuming task. In the initial stages of the development, the procedure will undergo many changes and each step needs to be described in laboratory records, which will crystallise into laboratory procedure(s) or a set of SOP(s) along the test development process. During this period a historical dataset on the reference and control items will also be collected. This dataset will be used to define the critical and relevant end-parameters and the control and reference items acceptance criteria.

Once the method is sufficiently developed and all parameters are defined, an in-house validation process during which the *in vitro* method is assessed for repeatability (accuracy & precision), selectivity, sensitivity, and stability should be performed (Section **8.3**). Likewise, its robustness is assessed (i.e., the influence of (external) parameters on the outcome parameters), as it is important to ensure the *in vitro* method performs in different laboratory environments, albeit within defined boundaries.

Figure 7.1. Evolution of a Standard Operating Procedure (SOP)



Optimisation of the SOP should be performed by following a formal procedure prior to formal in-house validation. It is critical that any parameter(s) to be optimised should be chosen prior to the optimisation process, should be measurable, so as to allow before and after comparison, and should include the optimisation steps to be performed. All data acquired during the optimisation steps should be annotated to allow tracking, comparison and measurement of the acquired optimisation.

During the in-house validation process (Section 8.3), weaknesses can come to light that may require adaptation or optimisation of the method, and which might also trigger the re-initiation of a new validation cycle. It is recommended that any intended changes introduced during the validation should be in the form of amendment(s) to the validation plan.

In addition, the *in vitro* method developer should be aware that if the *in vitro* method makes use of complex instrumentation and software, these commercial off-the-shelf products will require appropriate validation depending on the risk and the complexity of any customisation (OECD, 2016[2]). Spreadsheet templates using pre-defined formulas,

self-written equations, or macros developed for use with the *in vitro* method must also be validated and will also require documented procedures for correct use.

Upon a satisfactory completion of the validation process, the *in vitro* method development can be finalised and the final set of SOPs associated with the *in vitro* method will be available.

Once an *in vitro* method has been validated and published, e.g., in the format of an OECD Test Guideline (TG), the users will, from the published method, need to develop their own set of SOPs which are applicable and integrated into their organisation to assure the correct execution of the *in vitro* method within their facility's environment.

7.2. Preparing standard operating procedures

As indicated above, how to perform the *in vitro* method and related procedures is given in a set of SOPs, covering how to execute the *in vitro* method but also SOPs referring to general supporting procedures (e.g., test system handling, solubility assessment, cytotoxicity measurement, equipment maintenance, calibration and cleaning; handling of test and reference items; record keeping, reporting, storage, and retrieval, etc.). The reason for not having all *in vitro* method steps and processes described in one single SOP, but in a set of SOPs is for ease of use by the personnel involved and for facilitating regular review and updating.

The OECD GLP Principles (OECD, 1998[1]) provide examples of activities and processes that should be described in SOPs, while additional examples specific to *in vitro* testing are listed in the OECD advisory document on The Application of the Principles of GLP to *in vitro* Studies (OECD, 2004[3]).

SOPs should be written in the active voice and concisely explained in a step-by-step procedure, easy-to-read format. The information presented should be unambiguous and not overly complicated. The document should not be wordy, redundant, or overly lengthy but simple and short. The inclusion of a flow chart and/or a checklist to illustrate the process can help to make it clearer and more easily executable.

SOP(s) are best written by the individual(s) actually performing the work on a daily basis. The finalised SOP needs to be reviewed and approved by laboratory management, or the test facility manager in a GLP environment. SOPs are not static documents and need to be systematically reviewed on a periodic basis. SOP(s) may also need to be adapted whenever something changes (e.g., products, equipment and facility). In these cases a new version of the SOP should be approved. As soon as a new version is approved (date of approval), all concerned personnel need to be informed and the obsolete version removed from use and adequately archived. To allow and control this, all SOPs need to have a unique identifier (Title/version number/approval date). It is also recommended to detail the revision history in the document. In a GLP facility SOPs should be formally authorised by test facility management.

It is a good policy to divide SOPs by type, e.g., the EPA Guidance for Preparing Standard Operating details two types, technical SOPs and administrative SOPs. The guidance document provides examples and a standard layout for both technical and administrative SOPs (EPA, 2007[4]).

References

EPA (2007), *Guidance for Preparing Standard Operating Procedures*. [4]

OECD (2016), *Application of Good Laboratory Practice Principles to Computerised Systems*, OECD Series on Principles on Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing Paris. [2]

OECD (2004), *The Application of the Principles of GLP to in vitro Studies*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 14, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084971-en. [3]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [1]

Chapter 8. Performance of the method

Key message: *In vitro method developers need to ensure that in vitro methods they design will produce good quality data, i.e. fit for purpose, thanks to a stringent assessment of the performance of the method.*

Key content: *Elements of experimental design and how to determine the performance of a method are detailed, including aspects of plate layout, data analysis, and in-house method validation, including the assessment of linearity, range, accuracy, etc.*

Guidance for improved practice: *Details are given to increase the reliability of endpoint calculations when multiple independent experiments are run and to use tools to quantify performance characteristics.*

Recommendations *are given to in vitro method developers on how to increase the possibility of adoption of their method for regulatory purposes.*

In vitro method development and in-house validation should be considered as continuous and inter-dependent. Early in the development stage, the choice of instrumentation and methodology are selected based on the intended purpose and scope of the *in vitro* method. Once the development and optimisation of the method in the laboratory has been finalised it is recommended to perform an in-house validation (Section 8.3) of the method prior to routine use. This will provide documented evidence of the method performance in the laboratory and also prescribes on-going measures to ensure quality monitoring for the lifetime of the method. It will also check the feasibility of the method before the costly exercise of a formal collaborative trial (OECD , 2005[1]). The validation process, as described in the OECD Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment (OECD , 2005[1]), is not address this chapter (Section 1.3).

When discussing the in-house validation of *in vitro* methods, it is important to distinguish the analytical (measurable) endpoint (e.g. spectrophotometric, fluorimetric, mass-spectrometry and/or luminometric) from the *in vitro* method endpoint (e.g., proliferation, differentiation or viability) which refer to the biological concept being evaluated (Aschner et al., 2016[2]); (Schmidt et al., 2016[3]). The in-house validation of *in vitro* method analytical endpoint(s) is discussed in detail in (Section 8.3).

The assessment to be performed will largely depend on the type of *in vitro* method (i.e., qualitative or quantitative). Very few guidelines exist concerning the validation of qualitative methods (NATA, 2013[4]). There are several international guidelines available addressing the validation of quantitative methods (EMA, 2011[5]); (FDA, 2001[6]); (ICH, 2005[7]) which describe the general parameters required to assess the performance of the method analytical endpoint(s). International efforts have been made to discuss and harmonise the validation guidelines, however different interpretations still exist and no consensus regarding acceptance criteria or the actual validation process (Rozet *et al*., 2011) has yet been achieved. Most of these differences stem from the different regulatory frameworks in place. For a detailed comparison of the various international guidelines see (Kollipara et al., 2011[8]).

8.1. Acceptance criteria

As most *in vitro* methods are generally intended to predict a qualitative and/or a quantitative response, predictive of the degree of human or environmental hazard, it is essential that the *in vitro* method performs consistently over time and between laboratories.

Acceptance criteria should be developed based on historical data for all critical components and aspects of the method. Criteria should be defined for the test system (e.g., passage number, growth curve, cell recovery) and test system performance (e.g., positive, negative, and vehicle controls where applicable). Acceptance criteria should also be set for the analytical endpoint determination (e.g., linearity, accuracy, range) and also include data analysis (e.g., line fitting). These criteria should be developed and detailed in the *in vitro* method SOP(s).

Acceptance criteria should primarily be established based on information from historical data. When available, these can be then supplemented by data from validation studies, or from relevant bibliographic data including guidance documents. Historical data should be collected using the unchanged method, unless it can be shown that any changes have not affected the values. Data should only be rejected when there is a clear, valid and

scientifically justified reason to do so (Hayashi et al., 2011[9]), and the reasons for rejecting said data should be clearly and accurately documented.

For (transformed) data, which follows an approximate normal distribution, the mean and standard deviation (SD), e.g., for the positive control historical data, are calculated and the acceptance criteria are set at for instance ± 2 SD. For example the Bovine Corneal Opacity and Permeability (BCOP) *in vitro* method (Figure 6.1) uses 100% ethanol as the positive control. It has a mean published *in vitro* score (opacity + 15×permeability) of 51.6 ± 6.2 (mean ± Standard Deviation SD), which would set the acceptance criteria (mean ± 2 SD) to 39.2 to 64.0 (n=1171 trials) (Harbell and Raabe, 2014[10]).

For dose-response methods it is important to test multiple concentrations of the test item that fall within the linear dynamic range (Section 8.3.2) of the method, so as to narrowly define the 50% activity point. The 50% activity point (concentration) for the positive control may be used for e.g., establishing the acceptance criteria for a dilution-based cytotoxicity assay. This approach allows increased and decreased sensitivity to be readily identified.

Establishing acceptance criteria for the negative control is important to assure that the test system performs normally, and is just as important as for the positive control and can be done in the same manner (e.g., within +/- 2 or 3 SD of the historical mean response, or within the 95% control limits of the distribution of the historical data). Other acceptance criteria may be established such as criteria for the variability of the (quantifiable) data (e.g., OECD TG 431, 439, 492) or criteria for the minimum level of cell viability (e.g., OECD TG 442E).

Finally, it is also important to establish the cut-off value of the acceptance criteria, i.e., clear rules whether the response of a reference/control item is accepted or not, also taking into consideration the number of significant digits. The preferred approach is to specify the same number of significant figures both for the acceptance criteria and the measured result.

8.2. Experimental design

The number of replicates for each testing condition, including concentration level(s) used for the reference and control item(s), and test items etc., should be specified. During *in vitro* method development the number of replicates must be chosen using appropriate statistical methods. For example, a statistical power analysis (Crawley, 2015[11]) can be used to calculate the desirable number of replicates to detect a defined difference between treatments with pre-set levels of confidence (Krzywinski and Altman, 2013[12]). However, one should be aware that this number may be too high to be useful in practice. Alternatively the statistical power is provided for the chosen number of replicates.

Additionally, when multiple concentrations of a test item are tested, the mathematical model (e.g., dose-response curve) can be fitted to the experimental data using increasing number of tested concentrations and/or replicates, but generally it is better to increase the number of concentrations than the number of replicates. The lowest number of replicates that gives satisfactory variability of the parameter of interest (e.g., IC_{50} within acceptable limits) can be used in future studies (Assay Guidance Manual[1], High-Throughput Screening (HTS) Assay Validation[2]). Apart from these statistical considerations, sometimes practicalities such as cost and availability of replicates may also play a role in the selection process. However, the impact of reducing replicates should always be subjected to careful analysis and the corresponding statistical power should be given.

Similarly, the number of independent experiments needs to be evaluated. For instance, *in vitro* methods with a high degree of inter-experimental variability, such as those using primary tissues, may need a higher number of independent experiments compared to *in vitro* methods employing continuous cell lines.

Statistical methods (e.g., factorial design) can be very useful in the process of optimising new *in vitro* methods. To obtain an *in vitro* method that leads to accurate, reliable and robust readouts, the results of several combinations of any changes in the *in vitro* method would have to be assessed. Factorial design of experiments is often used where there are a large number of variables to be assessed, as it is nearly impossible to approach all possible combinations experimentally. It is efficient at evaluating the effects and possible interactions of several factors (independent variables). A statistical approach predicting the effect of changes in the *in vitro* method steps on the observed readout (known also as method robustness assessment) would allow for the development of an efficient *in vitro* method design, since the experimental robustness check can be based on a much smaller subset of combinations (Box, Hunter and Hunter, 2005[13]; Groten et al., 1997[14]).

8.2.1. Plate layout

When developing an *in vitro* method care should be taken to minimise any potential systematic effects. Many cell-based assays employ cell culture plates (6, 12, 96, or 384 well plates), so care must be taken to ensure cell seeding, treatment and measurement is performed in a uniform fashion across the whole plate (well-to-well), between plates and across multiple runs. Plate effects may occur in the outer wells (e.g., due to evaporation), across columns or rows or even within the actual wells (within well effects). Plating cell density is also crucial if exposure is taking place during the log phase such as for inhibition of cell growth endpoint. In this case if the control cells reach the stationary phase during the exposure time, the effect of the test item may be underestimated. Plate effects should be evaluated e.g., by using the same conditions/treatments across a complete plate.

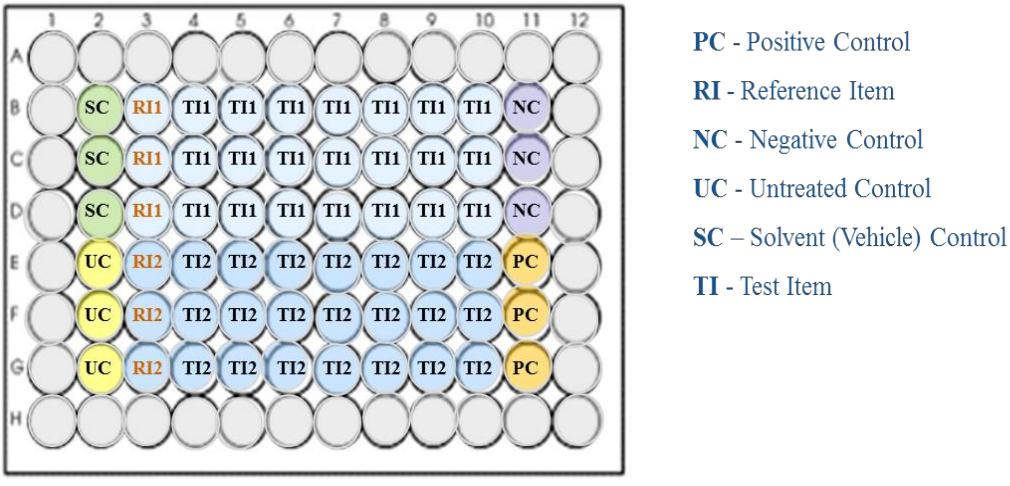
Drift can be due to seeding density variation during the process of initial cell seeding in plates, e.g., cells may be settling down in the master vessel which is used to store a cell suspension used to seed a particular plate. Additionally, using the same set of tips on a multichannel pipette while pipetting cells in media compositions prone to foaming, may compromise the accuracy of the seeding. Higher variability, which cannot be resolved via technique optimisation may require increased number of replicates/concentrations used to calculate the dose-response, or a higher numbers of independent experiments (Iversen et al., 2004[15]).

Randomisation of treatment wells in the cell culture plate is a strategy used to minimise inherent plate bias due to edge effects[3], drift, etc., and is particularly effective in an automated dosing setup. However randomisation may introduce other unforeseen errors, such as increased pipetting errors (usually only single wells can be pipetted), or data transfer/analysis errors (data may need to be rearranged for data analysis). It may also take significantly longer to treat the whole plate and so inadvertently introduce timing errors.

The plate layout will depend on the specific needs of the in vitro method, e.g., the numbers of controls, concentrations tested and replicates required. It should be such that cross-contamination (e.g., between test items) can be controlled for by checking variability between replicates. The plate design should also take into account how to perform comparison across plates so as to check between run or plate variability, by using

appropriate reference and control items. An example of an experimental 96-well plate layout using reference and control items is shown in Figure 8.1 (Coecke et al., 2014[16]).

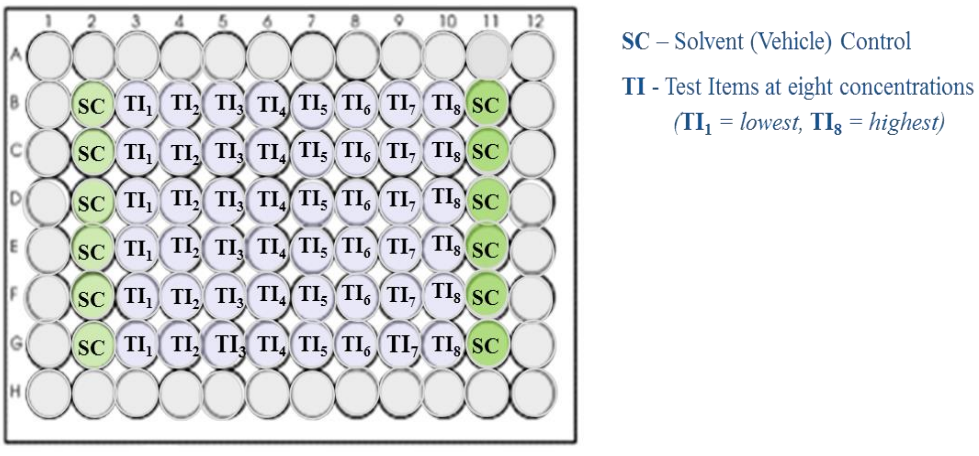
Figure 8.1. Example of plate layout including reference and control items, and solvent and untreated controls



Source: (Coecke et al., 2014[16])

The example plate layout minimises potential edge effects (difference between outer and inner wells due to evaporation). A way to assess plate drift is to include solvent controls (SC) on both the left and right side of the plate (Figure 8.2). Left and right SCs should not differ by more than a certain percentage for the plate to be accepted, e.g., a test meets acceptance criteria if the left and the right mean of the SCs do not differ by more than 15% from the mean of all SCs (NIH, 2001[17]).

Figure 8.2. Plate layout for systematic cell seeding errors



In addition, certain test or reference items may be volatile (e.g., solvents) or may contaminate neighbouring wells by capillary action, known as the wicking effect (Sullivan, 2001) and this may need to be taken into account in designing plate layouts. For instance, the commonly used cell lysis surfactant Triton X can affect cell viability in neighbouring wells and should be used at low concentrations or separated from cell-

containing wells by placing wells containing media or buffer in-between. Covering the plates with a foil prior to incubation, may also be employed, to avoid evaporation of volatile test items (e.g., OECD TG 442D).

Different effects found in the outer row of wells compared to the inner wells, are often due to uneven evaporation rates or plate stacking and can be a source of variation, as outer wells can often present as outliers compared to inner wells. Often the outermost wells contain a sterile, water-based solution and are not used for control, reference or test items as evaporation may take place during opening the door of the incubator. The cell-free wells may also be used for controls in the assay (e.g., background OD/FI, test item interference with the assay). Modern incubators are able to compensate much better for the change in humidity and so limit evaporation inside the incubator. If these outer wells are used for test, reference or control items, it should be clearly stated in the SOP to check for potential variation prior to use as not all laboratories may be equipped with appropriate incubators.

The inclusion of relevant reference and control items, and setting of acceptance criteria on the basis of historical data, is essential for regulatory applicability of *in vitro* methods and should be considered when developers decide on their plate layout. By including the correct reference and control items, the data set obtained from the *in vitro* method will demonstrate the correct functioning of the test system and the method used for analysis and therefore the validity of the experiments executed.

8.2.2. Data analysis

Transformation of data, e.g. normalisation or fitting to model equations, should be defined prior to data acquisition, and should be described in a SOP (OECD, 2017[18]) or in the relevant study plan. Formulas for normalisation (checked for accuracy) should be documented, validated (when implemented in electronic format) and disclosed along with a description and justification of the controls used in the calculation. It is recommended that computer scripts used to process raw data (e.g., Excel spreadsheets, scripts, macros etc.) should be validated and fully documented. The OECD Advisory Document Number 17 on the Application of GLP Principles to Computerised Systems provides guidance on validation of computerised systems (OECD, 2016[19]).

When a relationship is assumed between the tested concentrations and the response, a dose response curve, if required, can be fitted to obtain summary data such as the EC_{50} or IC_{50}. When fitting mathematical models, such as a dose-response curves or standard curves, to the data the models and reasoning behind their choice need to be documented. For example, when fitting a dose-response curve, the type of the equation used should be documented (e.g., a four parameter logistic curve) together with any constraints (e.g., top constrained to 100% in normalised data), limitations (e.g., assumption of monotonicity) and weightings (e.g., by inverse data uncertainty) applied (Motulsky, H.J.; Christopoulos, A., 2004[20]). Furthermore, the software name and version used to fit the equations should be documented, as well as the confidence interval of the parameters of interest, e.g., EC_{50}, RPC_{Max}, PC_{Max}, PC_{50} and/or PC_{10} (OECD, 2016[21]) and the relevant goodness of fit parameters (R-square, sum of squares etc.) so as to justify the selection of the model. In some cases it may be preferable to test multiple models (e.g., include non-monotonic curve) and select the best fit after the curve fitting.

8.2.3. Outlier detection and removal

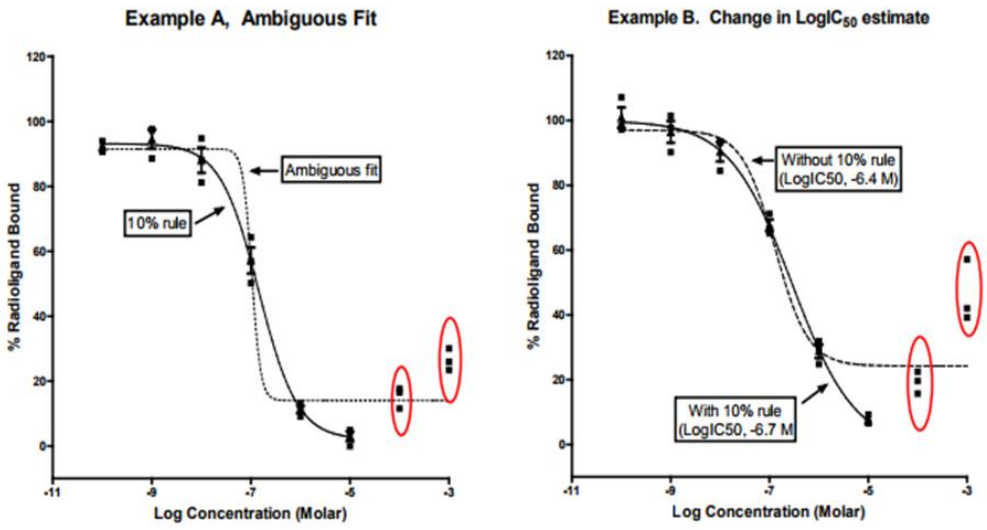
Criteria to detect/remove outliers should be stated and the reasoning should be given (Motulsky and Brown, 2006[22]; Pincus, 1995[23]). Outlier tests, such as Grubbs' Test (for single outlier) or using visual tools like boxplot (e.g., Mahalanobis distance), may be used to rule out outlier when it is difficult to judge whether the data should be regarded as an outlier or not, or when the reason for the occurrence of the outlier is unknown.

8.2.4. Non-monotonic dose and U-shaped curves

While dose response curves for relatively simplistic models, like most *in vitro* assays, are monotonic (i.e., they increase or decrease over the entire dose response range), sometimes a non-monotonic dose response can be observed. This potentially can be due to superimposition of separate effects that would individually elicit monotonic dose responses, or could indicate a perturbation of the system, e.g., due to chemical insolubility and/or precipitation at high concentrations (Section 6.5). Usually, such curves are U-shaped and can occur in many types of *in vitro* and binding assays.

An example of such a U-shaped curve is given below, where the analysis and interpretation of competitive binding data (Figure 8.3) can be complicated by an upturn of the percent binding when testing chemicals at the highest concentrations. As the concentration of the test item approaches the limit of solubility, the displacement of [^{3}H]17β-estradiol begins to generate a "U-shaped response curve" (OECD, 2015[24]). Such U-shaped curves are typically considered artefacts of the test conditions rather than relevant descriptors of the binding affinity of the test item. Retaining such data points (circled in red) when fitting competitive binding data to a sigmoid curve can inappropriately raise the perceived bottom of the curve Figure 8.3B, and can sometimes lead to a misclassification of the Estrogen Receptor (ER) binding potential for a test item (Figure 8.3A). This problem can be further controlled by excluding from the analyses all data points where the mean of the replicates for the % specific bound show 10% or more radioligand binding than the mean at a lower concentration (i.e., 10% rule), but it may not always be appropriate to include such a rule, as unintended and unforeseen consequences have been observed.

Figure 8.3. Example, Analysis of Competitive Binding Data, with and without use of 10% rule



If there is reason to believe that the test item is a non-binder, it might be appropriate to include a subjective waiver, so that the laboratories are allowed to use their judgement with regard to the use of the "10% rule", however this needs to be justified in the study report. It is important to note that a subjective waiver of the 10% rule may not be considered statistically appropriate.

8.3. In-house validation of the measurement endpoint(s)

This section focuses on the in-house validation of the measurement (analytical) endpoint(s), as described in the facility *in vitro* method SOPs. It is intended mainly for quantitative methods, however some of the performance characteristics relevant for qualitative methods are also described.

1. There are no specific regulations concerning *in vitro* method in-house validation, however guidance documents (e.g., FDA) and guidelines (e.g., EMA, ICH Q2(R1)) for analytical endpoint(s) validation describe the elements required in most validation studies in order to characterise the methods in terms of e.g., their reliability and reproducibility. The assessment to be performed will vary depending on whether the method has been developed in-house, whether it has been transferred from another laboratory, whether it is commercially available or whether it has previously undergone full validation. Many guidelines classify method validations into full validation, partial validation, and cross-validation (EMA, 2011[5]; FDA, 2001[6]).
2. A full validation includes all relevant aspects of a method, and should be performed for all new in-house developed methods or when major changes are made to an existing method.
3. A partial validation should be performed when a previously validated method undergoes minor modification(s), such as change in calibration concentration range or when published methods, often modified to the facility's requirements, are been transcribed into facility SOP(s).

4. A cross-validation is performed when two or more methods are used to generate data within the same study or across different studies. When performing cross-validation the same set of reference, control and/or test items should be analysed by both analytical methods.

While not all aspects of the in-house validation might be applicable to all methods, the guidance documents describe several acceptable approaches to be taken depending on e.g. the purpose of the method and whether the *in vitro* method is quantitative or qualitative. In general, quantitative methods should address at least, where applicable, the method's accuracy, reproducibility, linearity, limits of detection and range of measurement, while for qualitative methods specificity and sensitivity are key criteria.

Many of these guidance documents describe the principles of validation and when and how to apply them. Regardless of the type of validation (full, partial or cross-validation), it is recommended to develop a validation plan, preferably written as step-by-step instructions, where the appropriate performance characteristics to be assessed are defined up-front before beginning the validation. Some GLP concepts, such as the requirements for a study plan and a final report can be applied when performing an in-house validation study.

The validation should be performed by trained and qualified laboratory personnel, to minimise operator variability and only calibrated/validated (Section 4.1) equipment should be used, so as to reduce equipment related issues. The results of the validation should be compared with the acceptance criteria for the performance characteristics described in the validation plan and/or SOP(s) and any deviation should be recorded and documented in a summary report (validation report) together with conclusions on the outcome of the validation.

Care should be taken to select appropriate and meaningful performance characteristics, and not to miss potentially critical ones such as reagent variability when drafting the validation plan, as these will depend on the nature and type of the method being validated. A comparison of some of the quality parameters for both quantitative and qualitative methods is shown in Table 8.1.

Table 8.1. Common quality parameters for quantitative and qualitative analytical methods

Quantitative Method	Qualitative Method
Accuracy: trueness, precision	Sensitivity and specificity
Uncertainty	Unreliability region
Sensitivity and specificity	False positive and negative rates
Selectivity: interferences	Selectivity: interferences
Range and linearity	Cut-off limit
Detection limit	Detection limit
Ruggedness or robustness	Ruggedness or robustness

Source: (Trullols, Ruisánchez and Rius, 2004[25])

Caution must be applied when comparing these quality parameters as similar terms are used for different concepts and their evaluation may be different, e.g., for quantitative methods sensitivity should be a numerical value that indicates how the response changes whenever there is a variation in the concentration of the analyte. However for qualitative methods often reported as true/false or positive/negative, sensitivity is evaluated differently and as such may not be comparable (Section 8.3.4). The same applies to

specificity, detection limits, cut-off value and uncertainty or unreliability region (Trullols et al., 2005[26]) Some of the most common performance characteristics are discussed in the following sections.

8.3.1. Detection Limits and Cut-off values

The response of the instrument and the *in vitro* method with regard to the readouts of interest should be known, and should be evaluated over a specified concentration range, usually of the reference item. Various approaches may be used to determine the Limit of Detection (LOD) and Limit of Quantitation (LOQ) (ICH, 2005[7]).

1. Based on Visual Evaluation
2. Based on Signal-to-Noise ratio
3. Based on the Standard Deviation of the Blank[4]
4. Based on the Calibration Curve

Other approaches, described in the validation plan may also be employed. The LOD determines the lowest actual concentration or signal that can be consistently detected with acceptable precision, but not necessarily quantified. For normally distributed data, the LOD is often determined as the concentrations at the average response + 3 SD of the negative control range, as this gives only 1% chance of a false positive. LOQ is frequently calculated based on acceptable accuracy and precision of the reference item/reference item[5].

The Signal to Noise (S/N) ratio is frequently applied for methods which exhibit background noise (observed as the variation of the blanks) as baseline. It is calculated by comparing measured signals from samples with the reference item/positive control item with those of blank samples. A S/N ratio of 3 is generally accepted for estimating LOD, and S/N ratio of 10 is used for estimating LOQ (ICH, 2005[7]). Alternatively, assay acceptance can be determined using a signal window calculation[6].

For qualitative methods detection limits cannot be calculated as the SD can only be calculated when the response is a numerical value. For these methods a cut-off value, i.e., the minimum concentration of a substance needed to ascertain detection with a certain probability of error (usually 5%), can be calculated. The cut-off value is usually determined by establishing the false positive and negative rates at a number of levels below and above the expected cut-off concentration, and as such is related to the sensitivity of the method (Section 8.3.4).

8.3.2. Linearity and dynamic range

The response of the instrument detector can be expressed either as dynamic range or as linear dynamic range. The dynamic range is the ratio of the maximum and minimum concentration over which the measured property (e.g., absorbance) can be recorded. The linear dynamic range, i.e., the range of solute (e.g., reference item) concentrations over which detector response is linear, is more commonly used.

To quantify the amount of analyte in a sample a calibration curve is prepared, often assessed using a dilution series of the reference item. The results are plotted and a curve, usually linear, is fitted to the data. However not all *in vitro* methods will be linear for their full range, so a linear range (dynamic range) will need to be defined within the method's range.

For quantitative measurements, the boundaries of the dynamic range are determined by the lowest and highest analyte concentrations that generate results that are reliably produced by an *in vitro* method. The lower limit of linearity is frequently referred to as the Lower Limit of Quantification (LLOQ) and the upper limit of linearity as the Upper Limit of Quantification (ULOQ). The upper limit of linearity may be restricted by the highest available concentration in a sample or by the saturation of the signal generated by the instrument, while the lower limit is often limited by the instrument specifications.

The range is normally derived from the linearity and is established by confirming that the procedure provides an acceptable degree of linearity, accuracy and precision when applied to samples containing analyte (e.g., reference item) within or at the extremes of the specified range of the analytical method.

If a linear relationship exists statistical methods can be employed such as fitting of a regression line using the least squares and calculating the linear regression parameters (correlation coefficient, slope, y-intercept as well as residual sum of squares). Regression calculations on their own are usually considered insufficient to establish linearity and objective tests, such as goodness-of-fit may be required.

A correlation coefficient (r) of 0.99, based on a Goodness of Fit test, is often used as an acceptance criterion for linearity, however depending on the method lower r values may also be acceptable. Where a non-linear relationship exists, it may be necessary to perform a mathematical transformation of the data prior to the regression analysis. As linear regression is easy to implement, compared to other regression models (e.g., non-linear), a straight-line calibration curve will always be preferred.

For certain assays/methodologies, equations other than the linear can be fit as a standard curve, provided that the user is operating within the range of the assay/equipment (Section 4.1). However, it is recommended that the simplest model that adequately describes the concentration-response relationship is used. Selection of weighting and use of a complex regression equation should be justified. (Burd, 2010[27]; EMA, 2011[5]; FDA, 2001[6]; Viswanathan et al., 2007[28]). A minimum of 5 concentrations is recommended when assessing linearity, however other approaches may be used if justified (ICH, 2005[7]).

Subsequently, to facilitate efficient *in vitro* method transfer, the calculated linear regression parameters should be submitted along with a plot of the data. When the upper limit is exceeded (i.e., samples fall outside of the linear range), they may need to be diluted if possible. Where samples give a result below the lower limit of the linear range, it may be necessary to adapt the sample preparation to higher concentrations or change to a more sensitive apparatus.

8.3.3. Accuracy and precision

Assessment of accuracy and precision of a method will depend on whether the method is a quantitative or a qualitative method. The precision of a quantitative method describes the closeness of individual measures of an analyte (e.g. reference item) and is expressed as the coefficient of variation (CV). The FDA Bioanalytical Method Validation guidance document recommends that precision should be measured using a minimum of five determinations per concentration and a minimum of three concentrations in the range of expected concentrations (FDA, 2001[6]). Within-run and between-run precision should be reported.

For small molecules the within-run and between-run precision should not exceed 15% (20% at the LLOQ and ULOQ) while for large molecules (e.g. peptides and proteins) the within-run and between-run precision should not exceed 20% (25% at the LLOQ and ULOQ). The total error (i.e., sum of absolute value of the % relative error and % coefficient of variation) should not exceed 30% (40% at LLOQ and ULOQ) (EMA, 2011[5]).

For quantitative methods accuracy is usually determined using certified reference materials, if available, or by comparison to a reference method or to other methods. The accuracy of a method describes the closeness of mean test results obtained by the method to the actual value (or nominal) value (concentration) of the reference item. Accuracy is determined by replicate analysis of validation samples containing known amounts of the analyte (FDA, 2001[6]). The preparation of validation samples should mimic that of the study samples, and measurements should be made across at least 6 independent assay runs over several days (EMA, 2011[5]).

Accuracy should be reported as a percentage of the nominal value. When assessing the within-run and between-run accuracy for small molecules the mean concentration should be within 15% of the nominal value at each concentration level (20% at the LLOQ and ULOQ). For large molecules, both for within-run and between-run accuracy, the mean concentration should be within 20% of the nominal value at each concentration level (25% at the LLOQ and ULOQ) (EMA, 2011[5]).

Qualitative *in vitro* methods (e.g., as strong, weak), depend on accuracy and reliability to correctly classify chemicals according to its stated purpose (e.g., sensitivity, specificity, positive and negative predictivity, false positive and false negative rates). In such cases cut off values are used and their impact on the accuracy and reliability should be taken into account. The use of confidence bounds based on the distance from these cut-off values may not always be determined and therefore it may be preferable to conclude that the result is inconclusive (i.e., neither clear positive nor negative). The false positive rate is the probability that a test item which is actually negative being classified as positive by the method (Trullols et al., 2005[26]).

8.3.4. Sensitivity and specificity

For quantitative methods sensitivity may be defined as the capacity of the *in vitro* method to discriminate small differences in concentration or mass of the test item, while specificity may be defined as the ability of the *in vitro* method to identify, and where appropriate quantify, the analyte(s) of interest in the presence of other substances, i.e., the extent to which other substances may interfere with the identification/quantification of the analyte(s) of interest (ICH, 2005[7]). It may not be always possible to demonstrate that the method is specific for a particular analyte (complete discrimination). Specificity is concentration-dependent and is usually determined by adding materials which might be encountered in samples for which the method was developed. Specificity should be determined at the low end of the working range and ensure that the effects of impurities, cross-reacting substances, etc., are known.

Sensitivity in relation to qualitative methods may be defined as the ability of the method to detect the true positive rate while specificity is the ability of the method to correctly identify the true negative rate. The performance of a qualitative method can also be assessed with positive predictive values (PPV) and negative predictive values (NPV). PPV is the proportion of correct positive responses testing positive while NPV is the proportion of correct negative responses testing negative by an *in vitro* method

(Table 8.2). When calculating parameters such as sensitivity, specificity, false positive rate, false negative rate it is important that a balanced dataset is used (approximately an equal number of positive and negative compounds), otherwise these parameters will not reflect the true situation. The level of sensitivity, specificity, etc. which is acceptable is not standardised and is dependent on the list of items with which they are determined. Therefore, strict boundaries in acceptable levels for these accuracy parameters are not realistic. Generally though, sensitivities below 75% should not be accepted.

Table 8.2. Possible outcomes of an *in vitro* method result of a test item in a validation

		Condition		Prediction
		True	*False*	
Test Outcome	Positive	True positive (TP)	False positive (FP)	PPV
	Negative	False negative (FN)	True Negative (TN)	NPV
		↓ **Sensitivity**	↓ **Specificity**	**Accuracy**

$$\boldsymbol{Sensitivity}\ (\%) = \mathbf{100} \cdot \frac{\boldsymbol{TP}}{(\boldsymbol{TP} + \boldsymbol{FN})} \qquad \boldsymbol{Specificity}\ (\%) = 100 \cdot \frac{TP}{(TP + FN)}$$

$$\mathrm{PPV} = \frac{\text{number of true positives}}{\text{number of true positives} + \text{number of false positives}} \qquad \mathrm{NPV} = \frac{\text{number of true negatives}}{\text{number of true negatives} + \text{number of false negatives}}$$

8.3.5. Repeatability

Repeatability is defined as the closeness of the results between a series of measurements of a single sample obtained by the same study personnel, usually a single person, under the same operating conditions over a short interval of time, and is also called intra-assay precision. The most suitable means of expressing repeatability of an assay should be established following e.g., biostatistical evaluations. It is often expressed as % CV of a series of measurements, but may depend on the specific *in vitro* method and analytical methods being used.

8.4. Proficiency chemicals

For complex test methods, a number of proficiency chemicals are defined post-validation on the basis of the applicability domain and the dynamic range (i.e., spread of responses in the dataset) of the test method. These proficiency chemicals can be used by laboratories to demonstrate proficiency prior to the routine use of a validated test method or a test method falling within an adopted OECD test guideline. They usually represent either a subset of the reference chemicals included in the Performance Standards relating to the OECD TG, or chemicals used in the validation studies of the test method falling within the OECD TG.

Criteria used to select the proficiency chemicals for the *in vitro* method typically include chemicals that: i) represent the range of responses to be predicted, ii) have high quality reference data available; iii) cover the method's dynamic range of responses; iv) were correctly predicted by the test method during its validation study; v) cover a wide and representative range of relevant physical states, chemical classes, organic functional

groups and structures falling within the applicability domain of the in vitro method; vi) are commercially available and vii) are not associated with prohibitive acquisition and/or disposal costs.

It is also useful to include chemicals suspected of potentially interfering with the specific *in vitro* method format to better understand potential interference from unknown test items (examples include cytotoxic and cytostatic agents, fluorescent compounds and luciferase inhibitors and MTT interfering chemicals).

The number of proficiency chemicals will depend on the *in vitro* method type and purpose and should be chosen in such a way that a new laboratory can be confident that their results will be acceptable and robust. Since this greatly depends on the properties of the method, some methods may require 5 proficiency chemicals while for others up to 20 compounds should be tested.

In vitro method users should test the proficiency chemicals prior to routine testing for regulatory purposes and to formally comply with the OECD TG. The number of runs needed to correctly predict the proficiency chemicals is not limited. In this way, laboratories can demonstrate their proficiency in the *in vitro* method. Proficiency chemicals can also be used for training purposes, e.g., study personnel can demonstrate their ability to perform the method within the laboratory.

8.5. Data-intensive *in vitro* methods

The 21st century brought a paradigm shift in toxicity testing of chemical substances, relying more on higher throughput and/or high-content screening *in vitro* methods (National Research Council, 2007[29]). These allow the processing of hundreds or thousands of compounds simultaneously enabling the identification of mechanisms of action, and ultimately facilitating the development of predictive models for adverse health effects in humans. Furthermore, image analysis and omics-based *in vitro* method read-outs are getting more popular for *in vitro* method developers due to the data rich information obtained with such methods. The documentation and validation requirements for "data-intensive" approaches do not materially differ from those outlined in Section 8.3, but there may be additional specific aspects to address (e.g., the validity of the image-analysis approach used). It is recommended that the performance of high-throughput methods should be compared to "gold standard(s)", if available, or well-established methods (e.g., qPCR validation of key microarray/RNA-Seq findings). Further standardisation work will be required to achieve transferability and reproducibility of these *in vitro* methods.

The utility of "big data" for regulatory safety assessment has been addressed, e.g., omics (ECETOC, 2013[30]) or high throughput screening (Judson et al., 2013[31]). These data may be used in various contexts, such as supporting evidence for read-across, defining categories or to allow the design of Integrated Testing Strategies (ITS). Still, most applications have focused on screening and prioritisation as in the US EPA ToxCast program (Judson et al., 2010[32]; Judson et al., 2013[31]).

Although some technologies have been extensively used for decades (e.g., microarrays), debate is still ongoing about the interpretability and comparability of data generated from different sites and/or platforms. For many omics technologies consensus is still to be achieved concerning best practices in many critical aspects such as the experimental design and protocols for sample preparation and handling, data processing, statistical analysis and interpretation, and quality control (Bouhifd et al., 2015[33]). For some

technologies such as transcriptomics first respective frameworks have been proposed (Bridges et al., 2017[34]; Gant et al., 2017[35]; Kauffmann et al., 2017[36]).

The maintenance of high standards is essential for ensuring the reproducibility, reliability, acceptance, and proper application of the results generated. A certain level of standardisation is also needed since "big data" are generated using diverse technological platforms and various biochemical, analytical and computational methods, producing different data types and formats. Also to be addressed is the issue of "black box" validation for complex data processing routines.

Notes

1. See: https://www.ncbi.nlm.nih.gov/books/NBK53196/
2. See: https://www.ncbi.nlm.nih.gov/books/NBK83783/
3. See: http://labstats.net/articles/randomise_spatial.html
4. Blank: A sample of a biological matrix to which no analytes have been added that is used to assess the specificity of the bioanalytical method (FDA). An untreated control could be considered as a blank.
5. The reference item is often also used as a positive control
6. See: https://www.ncbi.nlm.nih.gov/books/NBK83783/#htsvalidation.Plate_Uniformity_and_Signa

References

Aschner, M. et al. (2016), "Upholding science in health, safety and environmental risk assessments and regulations", *Toxicology*, Vol. 371, pp. 12-16, http://dx.doi.org/10.1016/j.tox.2016.09.005. [2]

Bouhifd, M. et al. (2015), "Quality assurance of metabolomics", *ALTEX*, Vol. 32/4, pp. 319-326, http://dx.doi.org/10.14573/altex.1509161. [33]

Box, G., J. Hunter and W. Hunter (2005), *Introduction to fractional factorial experimentation*. [13]

Bridges, J. et al. (2017), "Framework for the quantitative weight-of-evidence analysis of 'omics data for regulatory purposes", *Regulatory Toxicology and Pharmacology*, Vol. 91, pp. S46-S60, http://dx.doi.org/10.1016/j.yrtph.2017.10.010. [34]

Burd, E. (2010), "Validation of Laboratory-Developed Molecular Assays for Infectious Diseases", *Clinical Microbiology Reviews*, Vol. 23/3, pp. 550-576, http://dx.doi.org/10.1128/cmr.00074-09. [27]

Coecke, S. et al. (2014), "Considerations in the Development of In Vitro Toxicity Testing Methods Intended for Regulatory Use", in *Methods in Pharmacology and Toxicology, In Vitro Toxicology Systems*, Springer New York, New York, NY, http://dx.doi.org/10.1007/978-1-4939-0521-8_25. [16]

Crawley, M. (2015), *Chapter 1. Fundamentals*, Wiley. [11]

ECETOC (2013), *Omics and Risk Assessment Science, Workshop Report No. 25*, European Centre for Ecotoxicology and Toxicology of Chemicals. [30]

EMA (2011), *Guideline on bioanalytical method validation*, European Medicines Agency, London. [5]

FDA (2001), *Guidance for Industry Bioanalytical Method Validation*, U.S. Department of Health and Human Services Food and Drug Administration, United-States. [6]

Gant, T. et al. (2017), "A generic Transcriptomics Reporting Framework (TRF) for 'omics data processing and analysis", *Regulatory Toxicology and Pharmacology*, Vol. 91, pp. S36-S45, http://dx.doi.org/10.1016/j.yrtph.2017.11.001. [35]

Groten, J. et al. (1997), "Subacute Toxicity of a Mixture of Nine Chemicals in Rats: Detecting Interactive Effects with a Fractionated Two-Level Factorial Design", *Fundamental and Applied Toxicology*, Vol. 36/1, pp. 15-29, http://dx.doi.org/10.1006/faat.1996.2281. [14]

Harbell, J. and H. Raabe (2014), "In Vitro Methods for the Prediction of Ocular and Dermal Toxicity", in *Handbook of Toxicology, Third Edition*, CRC Press, http://dx.doi.org/10.1201/b16632-6. [10]

Hayashi, M. et al. (2011), "Compilation and use of genetic toxicity historical control data", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 723/2, pp. 87-90, http://dx.doi.org/10.1016/j.mrgentox.2010.09.007. [9]

ICH (2005), *Validation of Analytical Procedures : Text and Methodology*. [7]

Iversen, P. et al. (2004), *Assay Guidance Manual*, Eli Lilly & Company and the National Center for Advancing Translational Sciences. [15]

Judson, R. et al. (2010), "In Vitro Screening of Environmental Chemicals for Targeted Testing Prioritization: The ToxCast Project", *Environmental Health Perspectives*, Vol. 118/4, pp. 485-492, http://dx.doi.org/10.1289/ehp.0901392. [32]

Judson, R. et al. (2013), *Perspectives on Validation of High-Throughput Assays Supporting 21st Century Toxicity Testing*, ALTEX, https://doi.org/10.14573/altex.2013.1.051. [31]

Kauffmann, H. et al. (2017), "Framework for the quality assurance of 'omics technologies considering GLP requirements", *Regulatory Toxicology and Pharmacology*, Vol. 91, pp. S27-S35, http://dx.doi.org/10.1016/j.yrtph.2017.10.007. [36]

Kollipara, S. et al. (2011), "International Guidelines for Bioanalytical Method Validation: A Comparison and Discussion on Current Scenario", *Chromatographia*, Vol. 73/3-4, pp. 201-217, http://dx.doi.org/10.1007/s10337-010-1869-2. [8]

Krzywinski, M. and N. Altman (2013), "Power and sample size", *Nature Methods*, Vol. 10/12, pp. 1139-1140, http://dx.doi.org/10.1038/nmeth.2738. [12]

Motulsky, H.J.; Christopoulos, A. (2004), *Fitting Models to Biological data using Linear and Nonlinear Regression: A Practical Guide to Curve Fitting*, Oxford University Press, Oxford (United Kingdom). [20]

Motulsky, H. and R. Brown (2006), , *BMC Bioinformatics*, Vol. 7/1, p. 123, http://dx.doi.org/10.1186/1471-2105-7-123. [22]

NATA (2013), *Guidelines for the validation and verification of quantitative and qualitative test methods*, National Association of Testing Authorities, Australia. [4]

National Research Council (2007), *Toxicity Testing in the 21st Century: A Vision and a Strategy*. [29]

NIH (2001), *Guidance Document on Using In Vitro Data to Estimate In Vivo Starting Doses for Acute Toxicity,*, National Institute of Health, United States. [17]

OECD (2017), *Guidance Document for Describing Non-Guideline In Vitro Test Methods*, OECD Series on Testing and Assessment, No. 211, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274730-en. [18]

OECD (2016), *Application of Good Laboratory Practice Principles to Computerised Systems*, OECD Series on Principles on Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing Paris. [19]

OECD (2016), *Test No. 455: Performance-Based Test Guideline for Stably Transfected Transactivation In Vitro Assays to Detect Estrogen Receptor Agonists and Antagonists*, OECD Guidelines for the Testing of Chemicals, Section 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264265295-en. [21]

OECD (2015), *Integrated Summary Report: Validation of Two Binding Assays Using Human Recombinant Estrogen Receptor Alpha (hrERa)*, OECD Publishing, Paris. [24]

OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, OECD Publishing, Paris.. [1]

Pincus, R. (1995), "Barnett, V., and Lewis T.: Outliers in Statistical Data. 3rd edition. J. Wiley & Sons 1994, XVII. 582 pp., £49.95", *Biometrical Journal*, Vol. 37/2, pp. 256-256, http://dx.doi.org/10.1002/bimj.4710370219. [23]

Schmidt, B. et al. (2016), "In vitro acute and developmental neurotoxicity screening: an overview of cellular platforms and high-throughput technical possibilities", *Archives of Toxicology*, Vol. 91/1, pp. 1-33, http://dx.doi.org/10.1007/s00204-016-1805-9. [3]

Trullols, E., I. Ruisánchez and F. Rius (2004), "Validation of qualitative analytical methods", *TrAC Trends in Analytical Chemistry*, Vol. 23/2, pp. 137-145, http://dx.doi.org/10.1016/s0165-9936(04)00201-8. [25]

Trullols, E. et al. (2005), "Validation of qualitative methods of analysis that use control samples", *TrAC Trends in Analytical Chemistry*, Vol. 24/6, pp. 516-524, http://dx.doi.org/10.1016/j.trac.2005.04.001. [26]

Viswanathan, C. et al. (2007), "Quantitative Bioanalytical Methods Validation and Implementation: Best Practices for Chromatographic and Ligand Binding Assays", *Pharmaceutical Research*, Vol. 24/10, pp. 1962-1973, http://dx.doi.org/10.1007/s11095-007-9291-7. [28]

Chapter 9. Reporting of results

Key message: *Good reporting of in vitro methods can only be achieved when all important details are recorded in a way that allows others to reproduce the work or reconstruct fully the in vitro method study.*

Key content: *Guidance is given on publishing and reporting of in vitro method studies and on data reporting for regulatory purposes.*

Guidance for improved practice: *Examples and available sources for scientific data management are detailed to promote more transparency and openness from scientists to avoid issues related to reproducibility of data but also to stimulate electronic data sharing for a variety of research and safety assessment purposes.*

Recommendations *are given to not only publish or make available the in vitro method results but also all the related documents and the changes that have been introduced to improve the method and the rational for them.*

In vitro methods must be fully documented following good recording and reporting practices, and must contain all pertinent details to allow subsequent and adequate analysis and reporting of results. For example, batch/lot numbers, catalogue numbers, supplier details, and expiry dates for chemicals and reagents must be listed for critical reagents, as well as temperatures and times (e.g., storage of chemicals, incubation steps in the *in vitro* method), specific identification of critical equipment used and, perhaps most importantly, any deviations (unintended variations) from Standard Operating Procedures (SOPs). All this information must be directly and accurately recorded, signed and dated by the person performing the activity, as these recordings are important for the correct interpretation of the results and reconstruction of the study. Where technical activities (e.g., aseptic work) preclude that person from recording the data themselves, the use of a second person to record the data may be employed. In these cases both the person performing the activity and the person recording the data must be identified in the study data.

Experimental details and results should be easily retrievable; a log page at the front of a notebook may help tracking recordings and observations. Any reference to computer files containing data should also be catalogued in the notebook. Data files should always be backed-up in case of computer failure, corruption, or deletion.

Reporting requirements depend on the different development phases of the *in vitro* method. For regulatory use, requirements for reporting are described in the Good Laboratory Practice (GLP) Principles. Reporting adequate information and results of all developmental phases will increase the confidence in the *in vitro* method and would allow for general acceptability by receiving authorities.

9.1. Publishing

There is an increasing tendency towards more transparency when publishing work which may lead to better reproducibility of published data, as described in the Guidelines for Transparency and Openness Promotion (TOP) Open Science Framework[1]. The EU Competitiveness Council has also announced as their target that all scientific publications resulting from publicly funded research should be publicly available by 2020[2]. It is also good practice to publish scientific results in a timely manner. The results will be used and re-used by other scientists, competitors, modellers or validation study statisticians. Moreover, for any systemic endpoint the prediction is/will be based on the results of many different studies, using different methods performed in different facilities, e.g. studies using *in vitro* methods for the identification of modulators of the thyroid hormone signalling pathway[3].

Sharing of data in public repositories is also being encouraged and best principles regarding the publication of scientific data have also been addressed by others, such as the FAIR (Findable, Accessible, Interoperable, and Reusable) Guiding Principles for scientific data management and stewardship, by the Nature Publishing Group[4]. This initiative not only promotes more transparency and openness but also promotes the use of computer readable datasets and data mining so that computers have the ability to access the data autonomously, unaided by their human operators, which is core to the FAIR Principles.

Data sharing is encouraged by default, unless there is reason for confidentiality, using public data-sharing standards and repositories such as ISA-TAB[5], Dryad Digital Repository[6], Figshare[7], and Nature Scientific Data[8]. It is recommended to not only publish the results, but also the method/SOP, again using on-line repositories such as

Nature Protocols[9], the Journal of Visualized Experiments (JoVE)[10], Testing method Exchange, Springer Protocols[11], EURL ECVAM Database service on Alternative Methods (DB-ALM)[12] and JRC-QSAR DB[13]. In the same vein, *in vitro* method modifications and further developments should be published. Such publications should include the changes leading to improvement, the rationale for them, and this should also entail information on which changes reduce *in vitro* method performance, or that do not result in an improvement.

In addition to the increasing openness and transparency, the publication of negative results is also gaining more ground e.g., the Journal of Negative Results in BioMedicine[14] is an open access, peer reviewed journal that provides a platform for the publication and discussion of non-confirmatory and "negative" data.

9.2. Mutual Acceptance of Data (MAD)

To avoid costly duplication in safety testing and government assessments, the OECD developed a framework for sharing data called the MAD system. MAD is a multilateral agreement which allows participating countries (including member states and MAD-adhering economies) to share the results of various non-clinical tests acquired when applying OECD methods and principles. As such it provides governments with confidence that non-clinical *in vitro* method data, generated under the MAD system, can be used in regulatory assessments. The use of this data by the Receiving Authority (Section 1.9) may differ depending on the scope of the specific test guideline, i.e., some *in vitro* methods may be full replacement, partial replacement, part of a defined approach or only used for screening purposes/priority setting. Results derived from non-standard *in vitro* methods and non-testing methods may also be reported, but as supporting information. Other benefits of the MAD system include reduction in animal testing, the evaluation of more chemicals and broader availability and transparency of government-vetted, high quality information and data.

9.3. Integrated Approaches to Testing and Assessment (IATA)

The current regulatory toxicity testing and assessment approach has evolved over the past half century, however it is unlikely to efficiently meet legislative mandates that require increased numbers of chemical assessments to be undertaken without a concomitant increase in the use of animals and resources. Therefore, new approaches are necessary to close the gap between the number of chemicals in use and the number assessed to date.

IATA[15] are pragmatic, science-based approaches for chemical hazard characterisation that integrates and weighs all relevant existing evidence and guides the targeted generation of new data, where required, to build up a hazard or risk assessment acceptable in regulatory decision-making. The information provided by individual *in vitro* methods, as well as *in silico* predictions, can be combined, interpreted and used for regulatory decision making by means of an IATA (OECD, 2017[1]). Ideally, an IATA should be informed by mechanistic understanding of the underlying toxicokinetics and toxicodynamics. A framework for capturing the toxicodynamic information is provided by Adverse Outcome Pathways (AOP)[16]. IATA and AOP knowledge, if properly captured and presented, leads to a better understanding of toxicity mechanisms, and ultimately the AOP knowledge derived from testing several chemicals may be extrapolated to predict the toxicity of all chemicals that trigger the same Molecular Initiating Event (MIE) or Key Event (KE).

Structured integration of different data types can be performed at different levels, including raw data and summarised level data (OECD, 2016[2]). Different levels of data integration can then be used including Boolean combinations of categorised results, scoring approaches, decision trees, deterministic and probabilistic approaches. As experience is gained, approaches to data integration can become standardised. Such approaches, called Defined Approaches (DAs), can thus become core elements of IATA. A DA is a formalised decision-making approach consisting of a fixed data interpretation procedure used to interpret data from a defined set of information elements (OECD, 2017[1]).

In contrast to IATA, DAs can be standardised and could therefore fall under MAD. The OECD is working on the development of a PBTG on DAs for skin sensitisation. The project, co-lead by the EC, US and Canada, aims at developing international standards that would give to DAs equal regulatory status as the current animal tests, i.e., prediction generated with valid defined approaches would fall under the OECD mutual acceptance of data program (Casati et al., 2017[3]).

It is essential to have all the results reported in a uniform manner to facilitate their use in the IATA framework, where the same dataset can be used in many different ways. The OECD GD 255 on reporting of DAs to be used within IATA provides a set of principles for reporting DAs to testing and assessment to facilitate their evaluation. Templates, for reporting individual and multiple information sources, are also available to provide consistent reporting which will ultimately facilitate the evaluation of IATA and DAs in regulatory decision-making within OECD Member Countries (OECD, 2017[1]).

9.4. Data reporting for regulatory purposes

Data and derived results from GLP studies will play an important role in increasing the relevance of *in vitro* data in regulatory contexts. Consideration and ultimately acceptance of *in vitro* GLP data can be promoted by using a standardised data format. This is facilitated by the use of IUCLID[17] (International Uniform ChemicaL Information Database), a software application used to record, store, maintain and exchange data on intrinsic and hazard properties of chemical substances.

The OECD had already designed and published several OECD Harmonised Templates (OHTs)[18] to report test results concerning:

1. physical/chemical properties (e.g., boiling point, density, flammability, etc.)
2. human toxicity (e.g., carcinogenicity, acute toxicity, etc.)
3. environmental toxicity (e.g., aquatic toxicity, terrestrial toxicity, etc.)
4. other properties describing degradation, accumulation etc.

These templates are geared towards results derived from classical (mostly OECD guideline) studies, focusing on apical endpoints, i.e., Adverse Outcomes (AOs).

However, reporting MIEs or KEs with such a classical OHT would tie them inseparably to the one AO the template covers, which is undesirable, as the *in vitro/in silico*/mechanistic information is then not easily accessible for building AOPs leading to other AOs: A Key Event can be relevant not only for one AOP, but several. Reporting the Intermediate Effect in an "AO-neutral" template makes the data available for all kinds of AOPs. A new, AO-neutral OHT was therefore needed that would allow reporting

observations from mechanistic (*in vitro* and *in silico*) tests, without immediately locking into one of several AOs the Intermediate Effect could lead to.

Knowing not only about results of animal tests (classical OHTs), but being able to cross-reference these test results with the intermediate effect observations (new OHT) has the potential to lead the way towards a less animal-centred hazard assessment. The OECD therefore started an initiative to come up with a stable, stakeholder-endorsed OHT for reporting on "intermediate effects" being observed via *in vitro* assays and possibly other non-animal test methods (computational predictions etc.). The template, OHT 201 - Intermediate effects, was endorsed by the OECD Joint Meeting in 2015 and was finally published in August 2016[19]. When submitting *in vitro* data to a receiving authority, the use of the OHT 201 is encouraged but is not yet obligatory.

The basic principle of OHT 201 is that:

1. one or several objective observation(s) (= results from non-classical test methods)
2. lead(s) to one subjective conclusion (= Intermediate Effect present, yes or no).

A properly filled in OHT 201 template therefore conveys a clear statement:

1. Based on observations O_1, O_2, …O_n
2. a certain chemical
3. triggers/does not trigger
4. a certain intermediate effect
5. on a certain biological level
6. at a certain effect concentration.

With OHT 201 being implemented in IUCLID[20], a software used by industry to fulfil reporting obligations under more and more legislative programmes (e.g., REACH), the concept of Intermediate Effects (and implicitly AOPs and predictive toxicology) has started to get attention in the regulatory world. This is a first step towards the acceptance of results from alternative tests for regulatory purposes, with the ultimate goal of replacing *in vivo* centred AO observations with alternative-methods-centred IATA/AOP considerations as the basis for risk assessment.

In the US if a chemical is not on the TSCA Chemical Substances Control Inventory[21], the substance is considered a "new chemical substance" while those already registered are considered as "existing chemical substances". Section 5 of TSCA requires anyone who plans to manufacture (including import) a new chemical substance for a non-exempt commercial purpose to notify the US EPA before initiating the activity. A pre-manufacture notice or PMN (a sample PMN form is available on the US EPA website[22]), must be submitted at least 90 days prior to the manufacture of the chemical.

For *in vitro* methods without a guideline, the Office of Pesticide Programs US EPA recommends following OECD Guidance Document 211 (OECD, 2017[4]) for describing non-guideline *in vitro* methods (EPA, 2016[5]).

9.5. Reporting of method validation

Validation is at the interface between *in vitro* method development/optimisation and regulatory acceptance/international recognition and ensures a science-based and conscientious evaluation of *in vitro* methods and approaches (e.g., Integrated Testing Strategies (ITS) or DAs), independent of specific interests, establishing their overall performance and fitness for a given purpose, i.e., their scientific validity[23].

The approach taken by a validation body may vary according to the needs of that body, e.g., whether they will coordinate the validation study or whether a validation study may be submitted to that body for assessment. In general an independent peer review of the validation study data and *in vitro* method is required, usually by organisations that specialise in *in vitro* method evaluations, such as JaCVAM[24], EURL ECVAM[25] or ICCVAM[26].

OECD published criteria should be met prior to seeking regulatory acceptance e.g., test method and validation study data should have been subjected to a transparent and independent peer review process, the generated data must be useful for hazard/risk assessment purposes, the submitted test method and data should adequately cover a spectrum of chemicals and products representative of those overseen by the receiving authority for which the method is proposed, the applicability and limitations of the test method should be clearly described and the test method should be time and cost effective and likely to be used in a regulatory context (OECD, 2005[6]). It is preferred that validation studies are performed and reported in accordance with the OECD Principles of GLP (OECD, 1998[7]). This will depend, however, on whether validation studies are part of the individual MA's inspection programme, as consensus has not been reached on this topic.

Submission of a new test method considered ready for proposal as an OECD Test Guideline is done via the OECD Secretariat either through a member country or through its National Co-ordinator; through the European Commission (EC) (EU only); an industry association through the Business and Industry Advisory Committee (BIAC) to the OECD; invited experts via a National Co-ordinator.

Notes

1. See: https://osf.io/ud578/
2. See: http://english.eu2016.nl/documents/press-releases/2016/05/27/all-european-scientific-articles-to-be-freely-accessible-by-2020
3. See: http://dx.doi.org/10.1787/9789264274716-en
4. See: http://www.nature.com/articles/sdata201618
5. See: http://isa-tools.org/
6. See: http://datadryad.org/
7. See: https://figshare.com/
8. See: http://www.nature.com/sdata/
9. See: http://www.nature.com/nprot/index.html
10. See: https://www.jove.com/
11. See: http://www.springerprotocols.com/
12. See: https://ecvam-dbalm.jrc.ec.europa.eu/
13. See: http://qsardb.jrc.it/qmrf/
14. See: https://jnrbm.biomedcentral.com/
15. See: http://www.oecd.org/chemicalsafety/risk-assessment/iata-integrated-approaches-to-testing-and-assessment.htm
16. See: http://www.oecd.org/chemicalsafety/testing/adverse-outcome-pathways-molecular-screening-and-toxicogenomics.htm
17. See: https://iuclid.echa.europa.eu/
18. See: https://www.oecd.org/ehs/templates/
19. See: http://www.oecd.org/ehs/templates/harmonised-templates-intermediate-effects.htm
20. See: https://iuclid6.echa.europa.eu/
21. See: https://www.epa.gov/tsca-inventory
22. See: https://www.epa.gov/sites/production/files/2017-02/documents/pmnviewonly.pdf
23. See: https://ec.europa.eu/jrc/en/eurl/ecvam/alternative-methods-toxicity-testing/validation
24. See: http://www.jacvam.jp/en/
25. See: https://eurl-ecvam.jrc.ec.europa.eu/
26. See: https://ntp.niehs.nih.gov/pubhealth/evalatm/resources-for-test-method-developers/submissions/index.html

References

Casati, S. et al. (2017), "Standardisation of defined approaches for skin sensitisation testing to support regulatory use and international adoption: position of the International Cooperation on Alternative Test Methods", *Archives of Toxicology*, Vol. 92/2, pp. 611-617, http://dx.doi.org/10.1007/s00204-017-2097-4. [3]

EPA (2016), *Process for Evaluating & Implementing Alternative Approaches To Traditional in Vivo Acute Toxicity Studies for Fifra Regulatory Use*, U.S. Environmental Protection Agency. [5]

OECD (2017), *Guidance Document for Describing Non-Guideline In Vitro Test Methods*, OECD Series on Testing and Assessment, No. 211, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274730-en. [4]

OECD (2017), *Guidance Document on the Reporting of Defined Approaches to be Used Within Integrated Approaches to Testing and Assessment*, OECD Series on Testing and Assessment, No. 255, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274822-en. [1]

OECD (2016), *Guidance Document for the Use of Adverse Outcome Pathways in Developing IATA for Skin Sensitisation*, OECD Series on Testing and Assessment, No. 260, OECD Publishing, Paris. [2]

OECD (2005), *Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, OECD Series on Testing and Assessment, No. 34, OECD Publishing, Paris. [6]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [7]

Chapter 10. Storage and retention of records and materials

Key message: *Before collecting data from in vitro methods it is important to assess the format of collection, the complexity involved and requirements for traceability, storage, verification and transmission of data.*

Key content: *The chapter gives insight on what key records and materials to archive and retain. It also details adequate document and record management of processes and the traceability of origin of materials.*

Guidance for improved practice: *Data integrity arrangements must be in place and structured methods and essential process components are described for both paper-based and electronic data to ensure that the collected in vitro method data are attributable, legible, contemporaneous, original and accurate.*

Recommendations *are detailed for the necessary procedures related to retention, archiving, retrieval, backup and restoration for all target groups involved in the in vitro method lifecycle.*

It is imperative that the historical data, paper-based or in the form of electronic data, are effectively managed so as to prevent any data integrity issues as this data may be requested when submitting the method for formal validation.

As compliance with the Principles of Good Laboratory Practice (GLP) is required by law for non-clinical safety studies in many (OECD) countries, it is important that newly developed *in vitro* methods are suitable to be performed in a GLP environment, and so avoid lengthy adaptation where possible (Coecke et al., 2016[1]). Studies which support validations may or may not be subject to verification depending on compliance monitoring authorities' programmes[1].

As the ultimate goal is to develop an *in vitro* method which will be formally validated for its future use in a regulatory environment following a quality system (e.g., GLP), it is essential to have some knowledge of the regulatory requirements specifically relating to the storage and retention of data, records and materials as the *in vitro* method should be designed so as to be easily transferrable into a GLP facility. In the early stages of method development there are less formal requirements for storage and retention of records and materials than in the later stages and in general facilities will follow internal policies regarding storage and retention of data, records or materials. The development phase should be used to define the raw data, preferably described in the *in vitro* method itself, and any data (e.g., metadata), records or materials, to be retained when used in a regulatory environment.

Before beginning to collect raw data from *in vitro* test procedures, it is important to assess the format of collection, the complexity involved and requirements for traceability, storage, verification and transmission of data. Specific standards may apply for data from regulatory testing and manufacturing (Coecke et al., 2005[2]); (FDA, 2003[3]); (OECD, 1999[4]). Data from material provided by tissue donors may also be subject to the requirements of data management and control under local, regional, national or international rules and regulations such as the EU Directive on Data Protection[2] (national and regional rules should be consulted as these may vary). It should be ensured that data reported accurately reflects the results obtained during experimental work, by performing adequate quality control of the data.

GLP test facilities should comply with the GLP principles with regards to storage and retrieval of records and data. The use of computerised systems and the generation of electronic data are now common across all aspects of a GLP study, through planning, performing, monitoring, recording and finally archiving. GLP data integrity requirements apply equally to paper and electronic data, and staff should be trained in data handling and data integrity and specifically with regards to ensuring electronic data integrity (Section 10.1).

Data may be generated in many ways, by recording manual observation, by printouts of simple equipment (e.g., balance) or by data generated using complex computerised systems. The more complex and configurable the system, the higher the risks to data integrity, however systems with lower complexity should not be overlooked. For instance, it may be relatively trivial to perform repeat measurements until the "correct" result is obtained (e.g., pH measurements). It is also important that all data is retained and archived.

Many electronic records are important to retain in their dynamic (electronic) format, to enable interaction with the data. Data must be retained in a dynamic form where this is critical to its integrity or later verification. This should be justified based on risk.

Stored data should be secured by both physical and electronic means against loss, damage and/or alteration. Stored data should be verified for restorability, accessibility, readability and accuracy. Verification procedures of stored data should be risk-based. Access to stored data should be ensured throughout the retention period (OECD, 2016[5]).

10.1. Data integrity

Data integrity arrangements must be in place throughout the *in vitro* method lifecycle to ensure that the accuracy and completeness of the data. The lifecycle includes all phases in the life of the data, records and materials, from their initial creation or purchase through processing, use, retention, archival and retrieval, and eventual destruction (if applicable). It is vital that formal records used to confirm the results and how they were obtained are held in a stable/secure form, duplicated (i.e., backed-up) and location which is documented and traceable and for which there is a minimum storage period. Disposal after such storage periods should be recorded and a summary report of the destroyed data and the means of destruction should be prepared and held.

If data is translated between different recording methods, systems and/or databases and, in particular critical phases like manual or semi-automatic transfer (e.g., Excel™ files to database, combination of information obtained from two or three databases to one database), correct resolution of pre- and post-translation data should be reviewed and confirmed by a qualified person. For handwritten data translated into an Excel™ sheet, it is also advisable for the changes to be verified by the same person who has made the observations. These issues are of special concern where data are exchanged between countries. When data translation occurs between different software or database systems, their compatibility and inability to be altered in translation should be tested and will need to involve appropriate validation procedures.

The acronym ALCOA (Attributable, Legible, Contemporaneous, Original, Accurate) has been widely associated with data integrity (FDA, 2007[6]); (WHO, 2016[7]). The Good Automated Manufacturing Practice (GAMP) guide "A Risk-Based Approach to GxP Compliant Laboratory Computerized Systems" includes an appendix (Appendix 3) on data integrity, which used the term "ALCOA +" described as **A**ttributable, **L**egible, **C**ontemporaneous, **O**riginal, **A**ccurate, complete, consistent, enduring, and available (Table 10.1).

Table 10.1. Terms associated with ALCOA

	Criteria	Description / Explanation	Comments
A	Attributable	Who performed an action and when? If a record is changed, who did it and why? Link to the source data.	**Who did it? Source data**
L	Legible	Data must be recorded permanently in a durable medium and be readable.	**Can you read it? Permanently recorded**
C	Contemporaneous	The data should be recorded at the time the work is performed and date / time stamps should follow in order.	**Was it done in "real time"?**
O	Original	Is the information the original record or a certified true copy?	**Is it original or true copy?**
A	Accurate	No errors or editing performed without documented amendments.	**Is it accurate?**

Comparisons are often made between secure electronic data and data that are available in paper format. The comparison results in similar conclusions that electronic data are more secure, more difficult to manipulate or change, and any changes are easier to detect

(assuming that the software is technically compliant to 21 Code of Federal Regulations (CFR) Part 11 and technical controls are appropriately implemented). On the other hand, changes to paper data, such as a printed chromatogram, are simpler to make, but may be much harder to detect.

The same principles should be applied when using either paper-based and electronic systems, or a combination of both. It should be assured that the data is unchanged from the source, and has not been modified, altered or destroyed. To ensure data integrity for both systems, the following components of this process should be taken into consideration.

1. **Documentation and result reporting**
   - Records must be clear and accurate.
   - All activities should be recorded at the time they are performed.
   - Records should also be chronological, traceable, and readily retrievable.
   - Original documents must be clearly identifiable (e.g., time stamps, watermarks) and standardised, predefined; authorised forms and templates should be used wherever possible and applicable.
   - Records should be signed and dated allowing for clear identification. The use of pencil either for recording data or signing/dating records should not be allowed. Recording of original (raw) data on loose notes or scrap sheets of paper should not occur. For electronic data, an audit trail recording who does what and when should be implemented.
   - Any corrections written on documents should be signed, dated and justified (i.e., indicate the reason for change) by a trained staff member, and must not obscure the original data.
   - Transcriptions, if performed need to be attached to the original results (full traceability) and reviewed.
   - Chronology of recorded data must be ensured.

2. **Effective review and verification**
   - A clear definition and understanding of raw data should be ensured.
   - There needs to be traceability to the testing method used, source data and verification of raw data.
   - SOPs need to be in place for data handling, record retention and good documentation practices and deviation handling etc.

3. **Additional considerations for electronic data**
   - If a system is required to maintain electronic data, it should be managed by unique user identity and password combination. If the system does not permit this, a paper-based log must be in place to record who uses the generic user and password combination, or who uses the unprotected equipment.

- Paper records can be reviewed for any changes or crossings out/deletions plus the signature/date and the reason for doing so. This is to be replicated in an electronic system in the same way by use of an electronic log (audit trail).
- Electronic records must be traceable to the operator who produced the records. Where there are multiple users, each user should be provided with a unique username/password combination and shared logins should not be allowed.
- There must be a periodic user account review procedure.
- There should be procedures in place for assigning access rights to each user.
- The level of access should be in line with the tasks that have to be performed.

4. Data storage

- Data must be stored in a safe and secure place for paper-based systems and in protected folders for electronic systems.
- An approach must be in place to ensure that data are protected against loss, damage or overwriting.
- Access to stored paper or electronic records must be restricted and tightly controlled and documented, e.g., original electronic data files may be saved as read-only, so as to avoid manipulation or loss of these files.
- Electronic records must be held in a format that is not readily corruptible and protected from deliberate or accidental alteration (e.g., CFR 21 part 11, GLP: see OECD GLP Guidance Document 17).

10.2. Retention and archiving

In a regulatory GLP environment the archiving retention time is sometimes defined in national legislation. However, where there is no retention time specified, the OECD recommends that records and materials should be retained for as long as receiving authorities might request GLP audits of the respective studies and at least three inspection cycles so that inspectors can evaluate the GLP compliance of the test facility and the respective studies (OECD, 2007[8]).

Retention arrangements must be designed to protect data, records and materials from deliberate or accidental changes, manipulations or deletions thus ensuring integrity throughout the retention period. Archiving is defined as the long-term retention of completed data and relevant metadata, records or materials. Archived data, records or materials may need to be stored for many years and must be permanently locked so that no changes can be made without detection. In the case of paper records, archive design and conditions must protect contents from untimely deterioration. In addition to this, they should be easily retrieved for regulatory inspections.

The archives must be designed so as to allow for the archiving of documents and records and also for the archiving of study samples and materials (e.g., slides, specimens, test items and reference material) under suitable storage conditions (OECD, 2007[8]). The OECD Principles of GLP state that: "a sample for analytical purposes from each batch of test item should be retained for all studies except short-term studies". The same rules

apply to these archives as apply to the paper based archive, i.e., access restrictions, retrieval and removal of items, etc.

The storage conditions should be optimal for these samples and often these archives will require dedicated storage facilities, e.g., low temperature storage such as -20°C, liquid nitrogen storage or storage of items under inert conditions. Where special storage equipment is required, the rules governing the control and maintenance of this equipment must be applied. Where computerised systems are used, these systems must also follow the facility's policy regarding the use of computerised systems, including qualification and validation of such systems (OECD, 2016[5]).

Samples of test and reference items or specimens may however be discarded when the quality of the material no longer permits evaluation. Obviously, the storage conditions should be optimal for these samples. It is also good practice to refer to the storage devices' history to determine equipment failures, power outages, moves, that could possibly impact sample integrity. When samples of test and reference items or specimens are disposed of before the end of the required retention period, the reason for disposal should be justified and documented (e.g., the reason might be perishable specimens such as blood smears, freeze-dried preparations and wet tissues).

Data is generated during the experimental phase of studies and during this phase the integrity of the data must be ensured until final archiving of the study. This data will usually be required for further analysis and as such will not be formally archived until the completion of the study. It is important that access to this data, both electronic and hard copies, is controlled until the final archiving upon completion of the study. It is recommended, where possible or feasible, that the electronic data is set as read-only or that an audit trail is provided, detailing who did what and when.

The GLP Principles for archiving must be applied consistently to electronic and non-electronic data. It is therefore important that electronic data is stored with the same levels of access control, indexing and expedient "retrieval" as non-electronic data. Electronic archiving should be regarded as an independent procedure which should be validated appropriately. A risk assessment should be applied when designing and validating the archiving procedure. Relevant hosting systems and data formats should be evaluated regarding accessibility, readability and influences on data integrity during the archiving period.

When electronic archiving is performed, the archiving system, both hardware and software, must be designated as a computerised system and validated as such so as to ensure the integrity of data stored electronically over its life-time. If the data storage media, the data formats, the hardware or software of the archiving system changes during the archiving period, the system should be revalidated so as to ensure that there is no negative influence on the accessibility, readability and integrity of the archived data and that the ability to retrieve the data has not been compromised. Where problems with long-term access to data are envisaged or when computerised systems have to be retired, procedures for ensuring continued readability of the data should be established. This may, for example, include producing hard copy printouts or converting data to a different format or transferring data to another system. If migration of data including conversion to a different data format or printing is relevant, the requirements of this guidance for data migration should be met. Risk assessment, change control, configuration management and testing regime should be considered as relevant standard procedures when changes in the archiving system are required. As content and integrity of any electronic data should be preserved during the archiving period, the complete information package should be

identified and archived (e.g., raw data, meta-data necessary to understand correctly the meaning of a record or to reconstruct its source, electronic signatures, audit trails, etc.).

10.2.1. Retrieval

Each facility should have in place procedures concerning the retrieval of archived records and materials. The procedures should detail who may retrieve records and materials, for how long and the return of records and materials to the archive. All steps mentioned above need to be documented and traceable.

In the case of electronic records, viewing the records without the possibility of alteration or deletion of the archived version does not constitute "retrieval" of a record. Most systems available nowadays support read-only access, without the possibility to change or delete the archived record.

10.3. Backup and restore

When storing electronic documents, including electronic archives, periodic backups should be performed. These backups do not constitute archived records, however as they may be required to be restored in the case a system failure, the same rules regarding access to the archived electronic records should be applied to access to the backup(s). In general backups are foreseen for short term storage and not long term storage or archiving and therefore the long-term readability of these archives is usually not an issue; however the restoration of the backups should also be checked on a regular basis.

Data generated during the experimental phase of the study should also be covered by the backup and restore policy of the facility.

Notes

1. See: http://www.oecd.org/env/ehs/testing/glp-frequently-asked-questions.htm
2. See: http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:31995L0046

References

Coecke, S. et al. (2005), "Guidance on good cell culture practice: A Report of the Second ECVAM Task Force on good cell culture practice", *ATLA Alternatives to Laboratory Animals*, Vol. 33/3, pp. 261-287. [2]

Coecke, S. et al. (2016), "Practical Aspects of Designing and Conducting Validation Studies Involving Multi-study Trials", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_5. [1]

FDA (2007), *Computerized Systems Used in Clinical Investigations*, Food and Drug Administration, United States. [6]

FDA (2003), *Part 11: Electronic Records; Electronic Signatures – Scope and Application*, Food and Drug Administration, United States. [3]

OECD (2016), *Application of Good Laboratory Practice Principles to Computerised Systems*, OECD Series on Principles on Good Laboratory Practice and Compliance Monitoring, No. 17, OECD Publishing Paris. [5]

OECD (2007), *Establishment and Control of Archives that Operate in Compliance with the Principles of GLP*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 15, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084995-en. [8]

OECD (1999), *Quality Assurance and GLP*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 4, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078598-en. [4]

WHO (2016), *Guidance on good data and record management practices*, WHO Technical Report Series, No. 996. [7]

Annex A. Good Cell Culture Practice (GCCP)

See the report:
https://www.researchgate.net/publication/7583114_Guidance_on_Good_Cell_Culture_Practice_-_A_report_of_the_second_ECVAM_task_force_on_Good_Cell_Culture_Practice

Annex B. Good Cell Culture Practice for stem cells and stem-cell-derived models

See the report : http://www.altex.ch/resources/altex_2017_1_095_132_t4_Pamies2.pdf

Annex C. EURL ECVAM GIVIMP meeting 24-25 February 2015 - participating experts

Maria Astridou	EFSA, Parma, Italy
Camilla Bernasconi	EURL ECVAM, Ispra, Italy
Gerard Bowe	EURL ECVAM, Ispra, Italy
Ann-Charlotte Bostroem	EURL ECVAM, Ispra, Italy
Sandra Coecke	EURL ECVAM, Ispra, Italy
Thomas Cole	EURL ECVAM, Ispra, Italy
Rita Cortvrindt	Vito, Mol, Belgium
Laureano Cuevas	Centro Nacional de Microbiologia, Majadahonda, Spain
Agnieszka Dudra	Bureau for Chemical Substances, Lodz, Poland
Eugene Elmore	University of California Irvine, Trabuco Canyon, Unites States
Chantra Eskes	SeCAM, Magliaso, Switzerland
Varvara Gouliarmou	EURL ECVAM, Ispra, Italy
Andrew Gray	United Kingdom GLP Monitoring Authority, The MHRA Medicines and Healthcare, London, United Kingdom
John Harbell	JHarbell Consulting LCC, Dallas, United States
Minne Heringa	RIVM, The National Institute for Public Health and the Environment, Centre for Safety of Substances and Products, Bilthoven, The Netherlands
Delyan Ivanov	University of Nottingham, Cancer Biology Queens Medical Centre, Nottingham, United Kingdom
Rob Jaspers	Health Care Inspectorate, The Netherlands
Ingrid Langezaal	EURL ECVAM, Ispra, Italy
Marcel Leist	University of Konstanz, Konstanz, Germany
Thomas Lucotte	ANSM, Inspectorate Department, Saint-Denis, France
Marisa Meloni	Vitroscreen SRL, Milano, Italy
Maurits-Jan Prinz	European Commission, Directorate General Internal Market, Industry, Entrepreneurship and SMEs, Brussels, Belgium
Vittorio Reina,	EURL ECVAM, Ispra, Italy
Glyn N Stacey	United Kingdom Stem Cell Bank, Advanced Therapies Division, NIBSC, United Kingdom
Emanuela Testai	Istituto Superiore di Sanita, Rome, Italy
Frederique van Acker	TNO Triskelion, Zeist, The Netherlands
Iwona Wilk-Zasadna	Novartis Vaccines, Basel, Switzerland.
Maria-Chiara Zorzoli	EURL ECVAM, Ispra, Italy

Annex D. OECD GIVIMP meeting 23-24 March, 2017 - participating experts

Luciene Balottin	Brazilian International Institute of Metrology, Quality and Technology (INMETRO), Brazil
Camilla Bernasconi	EURL ECVAM, Ispra, Italy
Ann-Charlotte Bostroem	EURL ECVAM, Ispra, Italy
Gerard Bowe	EURL ECVAM, Ispra, Italy
Patience Browne	OECD, Paris, France
Sandra Coecke	EURL ECVAM, Ispra, Italy
Maria Donner	Business And Industry Advisory Committee (BIAC)
Kristina Fant	SP Technical Research Institute of Sweden, Sweden
Detlef Hellmund	Ministry of Environment and Energy, Germany
Minne Heringa	RIVM, Netherlands
Miriam Jacobs	Department of Toxicology , Public Health England, United Kingdom
Rob Jaspers	Health Care Inspectorate, Ministry of Health, Welfare and Sports, Netherlands
Elisabeth Joossens	EURL ECVAM, Ispra, Italy
Yasunari Kanda	National Institute of Health Sciences, Japan
Emmanuel Lemazurier	INERIS - Direction des Risques Chroniques, France
Geoffrey Patton	Center for Food Safety and Applied Nutrition, Food and Drug Administration, United States
Emma Pedersen	SP Technical Research Institute of Sweden, Sweden
Markus Schulz	Business And Industry Advisory Committee (BIAC)
Tracey Holley	EURL ECVAM, Ispra, Italy
Sander Van Der Linden	EURL ECVAM, Ispra, Italy
Aneta Wiktorek-Smagur	Department of Good Laboratory Practices, Bureau for Chemical Substances, Poland

Annex E. Standardisation and accreditation bodies

Accreditation is a formal declaration by a neutral third party used to verify that laboratories have an appropriate quality management system and can properly perform certain test methods (e.g., ANSI, ASTM, and ISO test methods) and calibration parameters according to their scopes of accreditation. Organisations that issue credentials or certify third parties against official standards are themselves formally accredited by accreditation bodies, such as United Kingdom Accreditation Service (UKAS).

The International Organisation for Standardisation (ISO) is an independent, non-governmental membership organisation and the world's largest developer of International Standards with a central secretariat based in Geneva, Switzerland. The ISO story dates back to 1946 when delegates from 25 countries met at the Institute of Civil Engineers in London and decided to create a new international organization 'to facilitate the international coordination and unification of industrial standards'[1].

In this organisation, different industries define their specific technical standards and quality management requirements and issue ISO standards to guide conformity. ISO also standardises *in vitro* methods. Companies and organisations working according to ISO guidelines can ask for a conformity check and certification by independent accreditation bodies. ISO itself is not a controlling body, but has established a committee on conformity assessment (CASCO) guiding certification organisations.

While the OECD Principles of Good Laboratory Practice (GLP) and ISO/IEC 17025 both set out requirements for quality management systems under which testing is conducted, they have, as a result of their evolution and history, different purposes (OECD, 2016[1]).

ISO/IEC 17025 is an international standard intended to be applied by laboratory facilities conducting testing, according to established or specifically developed methodologies. The focus of the standard is on the on-going operation, monitoring and management of the laboratory itself, and on the capacity of the laboratory to produce consistent and reliable results that are scientifically valid. ISO/IEC 17025 can, in theory, be applied to any testing laboratory in any scientific discipline including those performing non-clinical testing. It is a reliable indicator of technical competence, and many industries routinely specify laboratory accreditation for suppliers of testing services.

Notes

1. See: http://www.iso.org

References

OECD (2016), *OECD Position Paper Regarding the Relationship between the OECD Principles of GLP and ISO/IEC 17025*, OECD Series on Principles of GLP and Compliance Monitoring, No. 18, OECD Publishing, Paris. [1]

Annex F. Good Laboratory Practice

Good Laboratory Practice (GLP) was developed in the 1970s in response to fraudulent scientific safety studies that were submitted to receiving authorities in support of applications for the regulatory registration/approval of drugs to the US FDA. Subsequently the principles of GLP were developed by the OECD to ensure data reliability and reconstructability of safety studies. The principles apply to all non-clinical health and environmental safety studies required by regulations for the purpose of registering or licensing chemical products of various kinds. The principles have been published in 1981 as an annex to the OECD Council Decision on Mutual Acceptance of Data (MAD)[1]. The decision states that 'data generated in the testing of chemicals in an OECD Member Country in accordance with the OECD Test Guidelines (Annex I of this decision) and OECD Principles of Good Laboratory Practice (Annex II of this decision) shall be accepted in other member countries for purposes of assessment and other uses relating to the protection of man and the environment'. Since 1981 a number of additional guidance, consensus and advisory documents have been published in the OECD Series on Principles of GLP[2], including an Advisory Document of the OECD Working Group on GLP n° 14 which specifically addresses *in vitro* Studies (OECD, 2004[1]).

In the EU, the principles of GLP are included in Directive 2004/10/EC, while the compliance monitoring procedures are included in Directive 2004/9/EC. GLP provisions are included in legislation for chemicals, human medicinal products, veterinary products, detergents, feed additives, food additives, genetically modified food or feed, pesticides, biocides and cosmetics (Coecke et al., 2016[2]), as well as for medical devices. Where applicable, conformity with the provisions of Directive 2004/10/EC of the European Parliament and of the Council shall be demonstrated. The US FDA requires GLP in the context of safety testing on medical devices.

Notes

1. See: http://www.oecd.org/env/ehs/mutualacceptanceofdatamad.htm
2. See: http://www.oecd.org/chemicalsafety/testing/oecdseriesonprinciplesofgoodlaboratorypracticeglpandcompliancemonitoring.htm

References

Coecke, S. et al. (2016), "Practical Aspects of Designing and Conducting Validation Studies Involving Multi-study Trials", in *Advances in Experimental Medicine and Biology, Validation of Alternative Methods for Toxicity Testing*, Springer International Publishing, Cham, http://dx.doi.org/10.1007/978-3-319-33826-2_5. [2]

OECD (2004), *The Application of the Principles of GLP to in vitro Studies*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 14, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264084971-en. [1]

Annex G. Solubility

How do we measure solubility?

• Solubility Determination
  ○ Visual inspection
  ○ Nephelometry
  ○ UV/VIS (absorbance)
  ○ UV/VIS following filtration step
  ○ Separation using High Performance Liquid Chromatography (HPLC) or Liquid Chromatography (LC) coupled with Mass Spectroscopic (MS) or UV detection (e.g., LC/MS, HPLC/MS)
• Several time points could be considered to make sure equilibrium is reached.

How does insolubility affect the concentration in an in vitro method?
• IC_{50} values can be shifted up if the test item precipitates as the effective concentration will be lower than the nominal concentration prepared. • Precipitates may also affect read-outs of the *in vitro* method and lead to impaired reproducibility
For some test items achieving suggested maximal target test concentration is difficult due to lack of solubility in test media.
• The exact threshold depends on the test item and the nature of the media used in the specific *in vitro* method. • The highest concentration in OECD TGs for mammalian cells for genotoxicity testing is 10 mM test item concentration in the tissue culture medium. Therefore, solubility needs to be achieved at a 1M in DMSO (so that the final concentration of DMSO on the cells is not higher than 1%). • There are exceptions which do not require full solubility e.g., OECD TG 442D, TG 442E and TG 487, where the highest test concentration is at the border of solubility, showing turbidity or slight precipitation (if not limited by other factors, such as cytotoxicity, pH, osmolality, etc.).
For these test items should there be guidance on how to establish the top concentration to test? For example if guidance indicates substances should be tested at top concentration of 1 mg/mL or 1mM but a substance does not go into solution at that concentration, is there some fractional concentration factor that

should be used to determine what next lower concentration to try?
• The spacing between concentrations is *in vitro* method-specific and can be used for solubility assessment as well. It is desirable to stay as close as possible to the precipitating concentration with the top concentration. • A preliminary test is often carried out to determine the appropriate concentration range of the test item to be tested, and to ascertain whether the test item may have any solubility and cytotoxicity issues (OECD TG 455) often using log-serial dilutions starting at the maximum acceptable concentration (e.g., 1 mM, 100 µM, 10 µM, maximum solubility, etc.) to find a concentration-response curve. Further runs, using smaller serial dilutions (e.g., 1:2, 1:3) are then used to focus in on the concentration-response curve, usually using six to eight concentrations (e.g., OECD TG 442E).
If it is difficult to get a test item in solution, at what point is the test item set aside as non-testable in the *in vitro* method?
• Test items are generally dissolved in a solvent (e.g., DMSO, ethanol, purified water). As a general rule, the final solvent concentrations should be as low as possible to avoid any potential interference with the *in vitro* method. Additional treatment(s) such as employing longer time frame, vortexing, sonication and/or heating may be if required. • In general test items should also be evaluated at low non-precipitating concentrations (if dissolved).
Is there a standard method or methods that could be used to accurately establish the solubility limits of test items so appropriate concentrations could be selected for testing?
• There are very accurate methods to determine the saturation point: e.g., analytical determination by HPLC and/or LC-MS/MS of concentration sampled from the supernatant, see OECD TG 105 for examples. • Nephelometric measurement of turbidity is much more accurate than visual evaluation (also possible in 96-well microtiter plates). • Precipitation can be identified with the eye quite easily. • It is important that the test facility has defined procedures (ideally SOPs) in place that describe how to conduct measurements and how to calibrate the procedure with known compounds depending on the intended applications. • Kinetic aspects should consider that there are compounds that need significant time to reach equilibrium.
Are there methods for determining solubility that work for some types of test items and not others?
• The choice of the solubility is dependent on the test item characteristics. UV/VIS methods must absorb light at the selected measurement wavelength to be employed. Nephelometry may suffer from interference when strong coloured items or items that fluoresce (e.g., contain a benzene ring) are

measured, e.g., phenol red which has light absorption in 430 and 560 nm and is excited by these wavelengths which results in fluorescence emission. • It is important to stay, when determining solubility, as close as possible to the real test conditions, where temperature and medium components such as pH, salts and proteins can influence solubility as most organic compounds absorb light in the UV range.
What are the set of acceptable solvents that are compatible with in vitro assays?
• As a common practice, organic solvents (e.g., DMSO, ethanol, methanol and acetone) are generally used to prepare the stock concentration even if the test item can also be dissolved in purified water. One of the reasons is that organic solvents prevent or minimise the growth of microorganisms which can then impact the test item stability over time. In case of a chemical/analytical method without living organisms/cells/tissues acetonitrile or methanol may be useful. • It is important to use a high purity of the solvent (95% to 100% purity). The final solvent concentration depends on the nature of the *in vitro* method but it needs to be less than 5% in most cases and can be as low as 0.1% [v/v].
If a test item is not soluble in the preferred solvent for a given *in vitro* method, how many of the other potential solvents should be tried? Will the compatible solvents be specific to each *in vitro* method and should the *in vitro* method define which are acceptable solvents and at what concentrations they are acceptable in the final test media?
• This is an *in vitro* method-specific question. Compatible solvents have to be defined by the *in vitro* method developer or user. And it has to be clearly demonstrated that the chosen concentration of the solvent has no adverse impact on the data. • There are recommendations published for specific *in vitro* methods (also in the OECD TG or related SOPs or scientific literature for new *in vitro* methods).
If a substance is highly soluble in an aqueous solution, can the substance be dissolved in the *in vitro* method media directly and tested without solvent carrier?
• Yes, it is the most preferred practise to test without adding any additional compound, so the absence of a solvent is highly welcome.

Annex H. Biokinetics and xenobiotic bioavailability

Since the techniques used for assessing biokinetics and xenobiotic bioavailability are complex, cost-intensive and time-consuming, routine use in the laboratory may not always be feasible, and therefore may prove more useful in troubleshooting the *in vitro* method.

Different processes result in a freely dissolved concentration that is not the same as the nominal concentration, (i.e., the added concentration). These processes are described in e.g., (Heringa et al., 2006[1]) and (Groothuis et al., 2015[2]), and were one of the main topic investigated by the FP7 EU Project Predict-IV, aimed to improve the predictivity of *in vitro* methods for unwanted effects of drugs after repeated dosing integrating biokinetics and biodynamic data. As one of the project outputs, a step-wise strategy was applied to measure and model cell exposure levels over time of a selected number of drugs in the developed *in vitro* assays. The strategy and the major obtained results are described in (Kramer et al., 2015[3]).

Evaporation / plastic and glass binding / sorption

In vitro systems are often open, with a small gap between the well plate and the lid, to allow air circulation for provision of oxygen for the cells and removal of excess CO_2. This air circulation allows volatile substances to evaporate into the air of the incubator. This may decrease the concentration in the medium in the test system, but can also contaminate medium in e.g., blank wells, as the substance can dissolve from the air into the medium of other wells present in the well plate or even the incubator. An example showing the effect of evaporation on test results can be found in Tanneberger (Tanneberger et al., 2010[4]). It may also be advisable to use tape/**foil** to cover culture plates in order to avoid evaporation of volatile substances and cross-contamination between wells **(e.g., OECD TG 442D).**

Lipophilic substances tend to bind to the plastic the cell culture plates are made of, although differences exist among the types of plastic used. The adsorption to polycarbonate is limited, but in organ-on-a-chip devices made of Polydimethylsiloxane (PDMS), there will be partitioning between the PDMS and the medium. PDMS is even used as an extraction material for Solid Phase Microextraction (SPME) (Heringa and Hermens, 2003[5]) and is therefore not suitable for *in vitro* test devices for testing of chemical substances. Examples where considerable binding to plastic was measured can be found in Kramer *et al.* (Kramer et al., 2012[6]) who also discuss how the addition of serum to medium decreases the binding to plastic. Other examples are reviewed in (Kramer et al., 2015[3]), reporting results of the Predict-IV project on cyclosporine A, amiodarone and chlorpromazine. The addition of serum to medium decreases the binding to plastic, but likely also the uptake into the cells (Pomponio et al., 2015[7]). Glass is a better material to avoid binding but very lipophilic substances are known also to bind to glass. Silanised glass can decrease this binding even further.

Sorption of the test item to cell-attachment matrices (e.g., collagen or matrigel layer used with hepatocytes in culture) is a specific aspect of interaction with the test device, although the relationship between a test item's lipophilicity and binding to is not as clear cut as it is for binding to plastic laboratory ware. The possible physical sequestration of test items, can lead to overestimating intracellular concentrations (Kramer et al., 2015[3]). Adsorption by coating material on plastics and feeder cells should also be considered.

Chemical degradation by hydrolyses and phototoxicity

The aqueous environment of the medium in an *in vitro* test enables spontaneous hydrolysis (i.e., without the aid of an enzyme) of substances with structures sensitive to this chemical reaction. During the time the test system, e.g., the well plate, is outside of the incubator, light will reach the medium and photolysis can take place for light sensitive substances. Therefore, information on hydrolysis and photolysis sensitivity is necessary before a substance is tested in an *in vitro* method (Section 6.2). More generally, each test facility should have adequate test item characterisation procedures in place to identify if the test item characteristics are compatible with the *in vitro* method.

Metabolism/metabolic stability

Some cell types have metabolic capacity, meaning that they contain significant levels of enzymes that convert the test substance to another substance. Especially cells originating from liver, intestine and lung are known to possess metabolic capacity, in decreasing order. In test systems with such cells, especially from these tissues, the concentration of the test item may decrease because of this metabolism, and the concentration of metabolites will increase. When a positive hazard response is obtained in such a cell system, it may thus either be caused by the test item itself, or its metabolite(s). Where there is a lag time in the response (compared to the positive control or other reference items), it could be that metabolite(s) are responsible (Pomponio et al., 2015[7]).

Protein binding

Protein binding can not only affect the freely dissolved concentration of substances as Heringa *et al* demonstrated where moderate differences in protein concentration in the test system resulted in the EC_{50} for a substance shifting by two to three orders of magnitude based on nominal concentration; but it also it is a factor to consider when comparing the responses between different *in vitro* systems. In systems where protein concentrations are relatively high, it requires more test item to achieve the same freely dissolved concentration, and therefore bioavailable concentration, as in those assays with lower protein concentrations. Testing at concentrations approaching the solubility limit of the test item within the test system does two things: first it provides the best experimental design to compare effects across *in vitro* test systems of similar endpoint but different protein concentration, and it reduces the potential for false negative results by testing to optimise the freely dissolved test chemical concentration in the test system. When working near the solubility limit of chemicals in test solution it can be helpful to include a concentration where solubility appears to be exceeded in the test run to better be able to distinguish it from concentrations where solubility is not exceeded. The next lowest test concentration would then be the concentration used to determine if there was activity in the assay for that test chemical. This strategy is particularly advantageous when the results of *in vitro* assays are to be used to identify chemicals with any potential for

activity, especially if further testing hinges on the results from these assays. Such a strategy has been considered to be advantageous in minimizing the chance of false negative results (Schmieder et al., 2014[8]).

The effect of protein in test systems is not only important for cell-based systems, but is also relevant to cell-free systems. For example, a receptor binding assay conducted using a cytosolic or nuclear preparation from a tissue may have more total protein than a competitive binding system using recombinant-expressed receptor protein. In the assay with greater concentration of non-receptor protein, it would take a higher nominal concentration of test item to displace the endogenous ligand from the receptor, and therefore a higher apparent IC_{50} based on nominal concentration when compared to a competitive binding assay with less total protein, but same receptor protein.

Serum is often added to cell culture medium to supplement it with important factors required for cell proliferation and maintenance. Serum-free medium is available and used, but not all cell types thrive in such culture conditions. Serum contains proteins, including albumin, which has non-specific binding sites, to which most organic substances tend to bind. As proteins are large molecules that do not transfer across a membrane, the binding to a protein renders a test item unavailable for cellular uptake, thus unable to reach any target inside the cell. Thus increasing the protein content in a test system can decrease the freely dissolved test item available to reach the target by shifting the equilibrium between freely dissolved and protein bound test item. Examples of the effect of serum protein binding can be found in (Heringa et al., 2004[9]) and in (Pomponio et al., 2015[7]).

On the other hand, serum proteins can also make some test items more accessible or more stable, e.g., for medical devices a medium with serum is preferred for extraction because of its ability to support cellular growth as well as to extract both polar and non-polar substances. In addition, protein binding also occurs *in vivo*. Therefore an *in vitro-in vivo* extrapolation method was developed to extrapolate nominal effective *in vitro* concentrations to equivalent *in vivo* plasma concentrations by accounting for the differences in protein concentrations (Gülden and Seibert, 2003[10]).

Cell membrane absorption

Cell membranes are composed of fatty acids, thus providing a lipid environment in which lipophilic substances will like to absorb. These absorbed molecules are then also not available for a target inside the cell. Examples showing the effect of membrane sorption can be found in Gülden *et al.* (Gülden, Mörchel and Seibert, 2001[11]) and in Bellwon *et al.*, (Bellwon et al., 2015[12]).

Measurement of free concentration/passive dosing

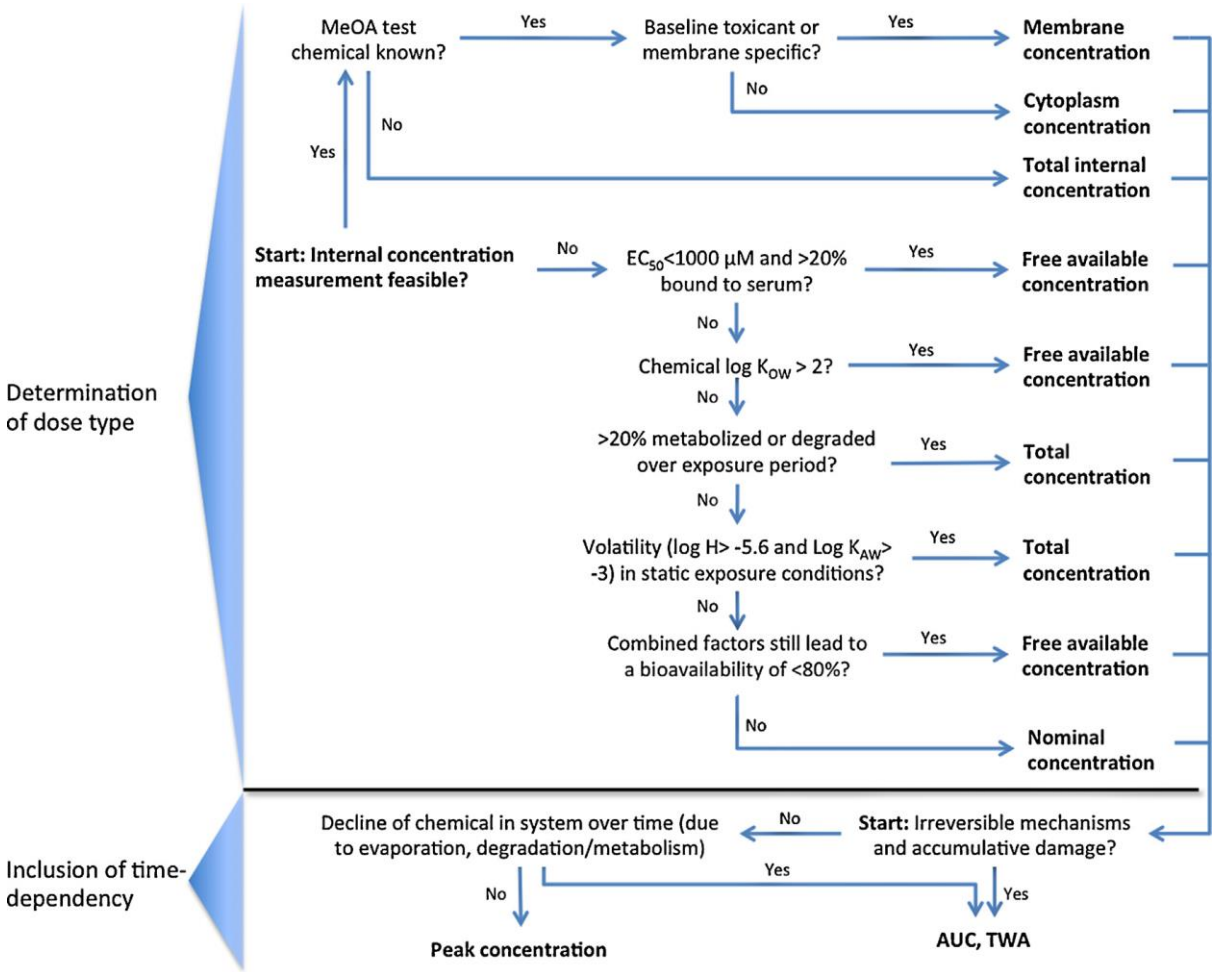
Clearly, several processes can influence how much of the test item will actually reach the target and is also related to its saturation concentration. If test results are based on the added, or nominal, concentrations, considerable variation between laboratories may be obtained. Furthermore, in vitro biokinetic processes are included in these test results (e.g., an EC50), rendering these unfit for extrapolation to in vivo (see In Vitro to In Vivo Extrapolation IVIVE below and (Kramer et al., 2015[3]). For example, if there is considerable evaporation, the EC50 in vitro will appear to be much higher than it will be in the same tissue in vivo. Thus, in order to obtain pure EC50 values, that relate target concentrations to responses, these target concentrations should be measured. As the precise concentration at the target site inside the cell is too difficult to measure, the best

approximation should be measured, i.e., the free concentration in the cell or in the medium. The free concentration in the cell is often still difficult to measure, therefore the free concentration in the medium (similar to the free concentration in the cell cytosol), or the total concentration in the cells (often for metals) are usually measured. Further information can then be added by calculations that take physicochemical and biochemical properties (e.g., transporter substrates) of the substances into account.

Methods with which the free concentration can be measured have been reviewed in Heringa *et al.* 2003 (Heringa and Hermens, 2003[5]). This review also describes how negligible depletion-solid phase extraction (nd-SPME) should be applied to measure free concentrations. This method is very suitable for *in vitro* tests, as it is suitable for small volumes. Examples of its application in *in vitro* tests are (Heringa et al., 2004[9]), (Broeders, Blaauboer and Hermens, 2011[13]) and (Kramer et al., 2012[6]).

Measuring the free concentrations does require extra effort and resources in the conduct of the *in vitro* test, as e.g., a chemical analysis method is necessary. This effort can be saved in some instances, depending on the properties of the test item: in case of very hydrophilic, non-volatile substances that hardly bind to serum proteins, there will hardly be any losses and the nominal concentration will be very similar to the free concentration. Figure A H.1 provides a decision scheme on which concentration should/can be used as dose metric (Groothuis et al., 2015[2]).

Figure A H.1. Flow chart to aid in choosing an appropriate dose metric for a specific *in vitro* toxicity test



Notes: First, a choice should be made for dose type based on the characteristics of the chemical and available knowledge. Then, the metric can be integrated or averaged in case of time-dependent exposure and irreversible mechanisms, or steady reduction over time. Peak concentration is defined here as the maximum concentration reached during the exposure period. Biokinetic/Toxicodynamic (BK/TD) may be applied to model partitioning and assess concentration changes over time.
Sources: The chart has been compiled by (Groothuis et al., 2015[2]) using literature data (Austin et al., 2002[14]); (Gülden et al., 2010[15]); (Gülden and Seibert, 2003[10]); (Knöbel et al., 2012[16]); (OECD, 2011[17]); (OECD, 2006[18]); (OECD, 2006[19]); (Reinert, Giddings and Judd, 2002[20]) and (Riedl and Altenburger, 2007[21]).

To avoid the effort of measuring free concentrations in every sample, passive dosing can be applied. In this method, a disk or ring of absorbent material, which is loaded with the test substance, is added to the sample. After a time of equilibration, the free concentration will have become proportionate to the concentration in the disk or ring, governed by the partition coefficient between water and the disk or ring material. If this partition coefficient has been predetermined, and if the amount of substance in the ring or disk by far exceeds the amount to be dissolved in the medium, then the free concentration in each sample can be easily calculated, and does not need to be measured. A more detailed description of the method (Smith, Oostingh and Mayer, 2010[22]) as well as a later study

(Smith et al., 2013[23]) provide examples of how passive dosing can be applied to *in vitro* tests.

In vitro to *in vivo* extrapolation (IVIVE) refers to the qualitative or quantitative transposition of experimental results or observations made *in vitro* to predict phenomena *in vivo*, on full living organisms. When the response of the *in vitro* test is plotted against the free concentration (or the nominal concentration only in case it can be demonstrated/estimated this approximates the free concentration), toxicity parameters such as the EC_{50} or a Benchmark Concentration (BMC) can be derived from the obtained curve. This *in vitro* toxicity parameter can be used as point of departure (PoD) for *in vitro* test circumstances and directly applicable to *in vivo* extrapolations (Blaauboer et al., 2012[24]); (Leist et al., 2014[25]) The corresponding *in vitro* concentrations can be converted into relevant plasma concentrations by taking the protein and lipid concentrations in plasma and cell culture medium into account (Bosgra and Westerhout, 2015[26]); (Zimmer et al., 2014[27]). In a final step, this concentration can be used as input for physiologically based pharmacokinetic (PBPK) models to estimate the dose that would result in the respective plasma concentration in man. This way an external Benchmark Dose (BMD) can be obtained. PBPK models describe the kinetic processes *in vivo*, relating external doses to tissue concentrations in time. For these models, some physical-chemical properties of the test substance need to be known, as well as some kinetic parameters such as the fraction absorbed, rate of metabolism, tissue partition coefficients, protein binding coefficients and urinary excretion rate (Louisse et al., 2010[28]). Good modelling practices for Physiologically Based PharmacoKinetic (PBPK) models have been described by Loizou *et al.* (Loizou et al., 2008[29]). The recommendations from a joint EPAA - EURL ECVAM on how Physiologically Based ToxicoKinetic (PBTK) modelling platforms and parameter estimation tools could enable animal-free risk assessment are reported in Bessems *et al.*, (Bessems et al., 2014[30]).

References

Austin, R. et al. (2002), "The Influence of Nonspecific Microsomal Binding on Apparent Intrinsic Clearance, and Its Prediction from Physicochemical Properties", *Drug Metabolism and Disposition*, Vol. 30/12, pp. 1497-1503, http://dx.doi.org/10.1124/dmd.30.12.1497. [14]

Bellwon, P. et al. (2015), "Kinetics and dynamics of cyclosporine A in three hepatic cell culture systems", *Toxicology in Vitro*, Vol. 30/1, pp. 62-78, http://dx.doi.org/10.1016/j.tiv.2015.07.016. [12]

Bessems, J. et al. (2014), "PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment", *Regulatory Toxicology and Pharmacology*, Vol. 68/1, pp. 119-139, http://dx.doi.org/10.1016/j.yrtph.2013.11.008. [30]

Blaauboer, B. et al. (2012), "The use of biomarkers of toxicity for integrating in vitro hazard estimates into risk assessment for humans", *ALTEX*, Vol. 29/4, pp. 411-425, http://dx.doi.org/10.14573/altex.2012.4.411. [24]

Bosgra, S. and J. Westerhout (2015), "Interpreting in vitro developmental toxicity test battery results: The consideration of toxicokinetics", *Reproductive Toxicology*, Vol. 55, pp. 73-80, http://dx.doi.org/10.1016/j.reprotox.2014.11.001. [26]

Broeders, J., B. Blaauboer and J. Hermens (2011), "Development of a negligible depletion-solid phase microextraction method to determine the free concentration of chlorpromazine in aqueous samples containing albumin", *Journal of Chromatography A*, Vol. 1218/47, pp. 8529-8535, http://dx.doi.org/10.1016/j.chroma.2011.09.064. [13]

Groothuis, F. et al. (2015), "Dose metric considerations in in vitro assays to improve quantitative in vitro–in vivo dose extrapolations", *Toxicology*, Vol. 332, pp. 30-40, http://dx.doi.org/10.1016/j.tox.2013.08.012. [2]

Gülden, M. et al. (2010), "Cytotoxic potency of H2O2 in cell cultures: Impact of cell concentration and exposure time", *Free Radical Biology and Medicine*, Vol. 49/8, pp. 1298-1305, http://dx.doi.org/10.1016/j.freeradbiomed.2010.07.015. [15]

Gülden, M., S. Mörchel and H. Seibert (2001), "Factors influencing nominal effective concentrations of chemical compounds in vitro: cell concentration", *Toxicology in Vitro*, Vol. 15/3, pp. 233-243, http://dx.doi.org/10.1016/s0887-2333(01)00008-x. [11]

Gülden, M. and H. Seibert (2003), "In vitro–in vivo extrapolation: estimation of human serum concentrations of chemicals equivalent to cytotoxic concentrations in vitro", *Toxicology*, Vol. 189/3, pp. 211-222, http://dx.doi.org/10.1016/s0300-483x(03)00146-x. [10]

Heringa, M. and J. Hermens (2003), "Measurement of free concentrations using negligible depletion-solid phase microextraction (nd-SPME)", *TrAC Trends in Analytical Chemistry*, Vol. 22/9, pp. 575-587, http://dx.doi.org/10.1016/s0165-9936(03)01006-9. [5]

Heringa, M. et al. (2006), "Measurement of the free concentration of octylphenol in biological samples with negligible depletion-solid phase microextraction (nd-SPME): Analysis of matrix effects", *Journal of Chromatography B*, Vol. 834/1-2, pp. 35-41, http://dx.doi.org/10.1016/j.jchromb.2006.02.009. [1]

Heringa, M. et al. (2004), "Toward More Useful In Vitro Toxicity Data with Measured Free Concentrations", *Environmental Science & Technology*, Vol. 38/23, pp. 6263-6270, http://dx.doi.org/10.1021/es049285w. [9]

Knöbel, M. et al. (2012), "Predicting Adult Fish Acute Lethality with the Zebrafish Embryo: Relevance of Test Duration, Endpoints, Compound Properties, and Exposure Concentration Analysis", *Environmental Science & Technology*, Vol. 46/17, pp. 9690-9700, http://dx.doi.org/10.1021/es301729q. [16]

Kramer, N. et al. (2015), "Biokinetics in repeated-dosing in vitro drug toxicity studies", *Toxicology in Vitro*, Vol. 30/1, pp. 217-224, http://dx.doi.org/10.1016/j.tiv.2015.09.005. [3]

Kramer, N. et al. (2012), "Quantifying Processes Determining the Free Concentration of Phenanthrene in Basal Cytotoxicity Assays", *Chemical Research in Toxicology*, Vol. 25/2, pp. 436-445, http://dx.doi.org/10.1021/tx200479k. [6]

Leist, M. et al. (2014), "Consensus report on the future of animal-free systemic toxicity testing", *ALTEX*, pp. 341-356, http://dx.doi.org/10.14573/altex.1406091. [25]

Loizou, G. et al. (2008), "Development of good modelling practice for physiologically based pharmacokinetic models for use in risk assessment: The first steps", *Regulatory Toxicology and Pharmacology*, Vol. 50/3, pp. 400-411, http://dx.doi.org/10.1016/j.yrtph.2008.01.011. [29]

Louisse, J. et al. (2010), "The Use of In Vitro Toxicity Data and Physiologically Based Kinetic Modeling to Predict Dose-Response Curves for In Vivo Developmental Toxicity of Glycol Ethers in Rat and Man", *Toxicological Sciences*, Vol. 118/2, pp. 470-484, http://dx.doi.org/10.1093/toxsci/kfq270. [28]

OECD (2011), *Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264069923-en. [17]

OECD (2006), *Draft proposal for a new guideline. In: Fish Embryo Toxicity (FET)*, OECD Publishing, Paris. [18]

OECD (2006), *Test No. 221: Lemna sp. Growth Inhibition Test*, OECD Guidelines for the Testing of Chemicals, Section 2, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264016194-en. [19]

Pomponio, G. et al. (2015), "In vitro kinetics of amiodarone and its major metabolite in two human liver cell models after acute and repeated treatments", *Toxicology in Vitro*, Vol. 30/1, pp. 36-51, http://dx.doi.org/10.1016/j.tiv.2014.12.012. [7]

Reinert, K., J. Giddings and L. Judd (2002), "Effects analysis of time-varying or repeated exposures in aquatic ecological risk assessment of agrochemicals", *Environmental Toxicology and Chemistry*, Vol. 21/9, pp. 1977-1992, http://dx.doi.org/10.1002/etc.5620210928. [20]

Riedl, J. and R. Altenburger (2007), "Physicochemical substance properties as indicators for unreliable exposure in microplate-based bioassays", *Chemosphere*, Vol. 67/11, pp. 2210-2220, http://dx.doi.org/10.1016/j.chemosphere.2006.12.022. [21]

Schmieder, P. et al. (2014), "A rule-based expert system for chemical prioritization using effects-based chemical categories", *SAR and QSAR in Environmental Research*, Vol. 25/4, pp. 253-287, http://dx.doi.org/10.1080/1062936x.2014.898691. [8]

Smith, K. et al. (2013), "The dosing determines mutagenicity of hydrophobic compounds in the Ames II assay with metabolic transformation: Passive dosing versus solvent spiking", *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, Vol. 750/1-2, pp. 12-18, http://dx.doi.org/10.1016/j.mrgentox.2012.07.006. [23]

Smith, K., G. Oostingh and P. Mayer (2010), "Passive Dosing for Producing Defined and Constant Exposure of Hydrophobic Organic Compounds during in Vitro Toxicity Tests", *Chemical Research in Toxicology*, Vol. 23/1, pp. 55-65, http://dx.doi.org/10.1021/tx900274j. [22]

Tanneberger, K. et al. (2010), "Effects of Solvents and Dosing Procedure on Chemical Toxicity in Cell-Basedin VitroAssays", *Environmental Science & Technology*, Vol. 44/12, pp. 4775-4781, http://dx.doi.org/10.1021/es100045y. [4]

Zimmer, B. et al. (2014), "Profiling of drugs and environmental chemicals for functional impairment of neural crest migration in a novel stem cell-based test battery", *Archives of Toxicology*, http://dx.doi.org/10.1007/s00204-014-1231-9. [27]

Annex I. List of viability testing methods (non-inclusive) of cell cultures

1. Structural cell damage (non-invasive)	
Evaluation of overall cell shape, cytoplasmic structure, flatness and outline properties on a good phase contrast light microscope	Screening assay that covers many forms of damage with high sensitivity, if observer is experienced. May be automated and rendered quantitative to some extent by high content imaging. • **Advantages:** high throughput (if automated), non-invasive, repeatable on same well over time. • **Disadvantages:** No clear prediction model (only qualitative data, no exact cell death definition), no standalone approach; requires extensive experience of operator.
LDH-release test	Cells with intact membrane retain their content of LDH enzyme; LDH is released when cell membranes rupture (non-viable cells), and the enzyme can then be measured in the supernatant. To give fully quantitative data, the assay requires normalisation to the total LDH content of the positive control well(s). It can to some extent be repeated for the same culture at different time points. • **Advantages:** Measurement of a definite/unambiguous cell death endpoint; can be combined with cell function assays. Allows cells to be used for other purposes, if only supernatant is sampled. • **Disadvantages:** Information only for cell populations. Requires normalisation to the total LDH content of a culture well (extra wells for cytotoxicity positive control treated). Frequently high background LDH levels are observed (e.g., from serum components; signal/noise ratio can be bad in some culture media or with some cell types). Problems with long-term assays involving medium changes. Not a very sensitive measure of cytotoxicity
2. Structural cell damage (invasive)	
Membrane penetration using dyes to detect 'cytotoxicity' (e.g., naphthalene black, trypan blue, propidium iodide, ethidium bromide, EH-1)	Dyes are selected so that they stain non-viable cells, but do not enter cells with an intact cell membrane. Some of the dyes stain the entire cell (e.g., trypan blue), others stain the nucleus/DNA (e.g., propidium iodide). Dyes that only stain dead cells usually need a combination with a method that stains/identifies all cells (such as phase contrast for trypan blue, or a nuclear counterstain (H-33342, acridine orange, SYTO-13) for fluorescent dyes. • **Advantages:** Rapid and usually easy to interpret. Gives information

	on the single cell level. High throughput and absolute quantification are possible (high content imaging). • **Disadvantages:** May overestimate viability since apoptotic cells continue to have intact membranes and may appear viable. Some dyes (e.g., trypan blue, H-33342) are cytotoxic, so that the evaluation has to be performed rapidly. Trypan blue and ethidium bromide, are toxic/CMR classified chemicals and the use should be restricted.
Retention of dyes within intact cells to detect 'viability' (e.g., fluorescein diacetate or calcein-AM)	After dye exposure, viable cells fluoresce when observed under UV light. The lipid-soluble dyes are transformed by cellular enzymes (esterases) into lipid-insoluble fluorescent compounds that cannot escape from cells with intact membranes. Thus, cells can be observed under a microscope (single cell analysis) or with a fluorescence plate reader (population analysis). The dyes are often used in combination with a cytotoxicity stain (e.g., propidium iodide). • **Advantages:** Rapid and usually easy to interpret. Gives information on the single cell level (including morphological information on the cell shape). High throughput and absolute quantification are possible (high content imaging, fluorescent plate reader or FACS). • **Disadvantages:** Some cells leak the dyes; some cells actively export the dyes through P-gp activity. Many fluorescent dyes are prone to photo-bleaching, and some may be sensitive to their local environment (pH etc.).
Evaluation of programmed cell death/apoptosis markers	As programmed cell death is a universal biological process based on defined cellular biochemical pathways and organelle changes, the activation of cell-death-associated pathways is often used as surrogate marker for cell death. An example for such a pathway is the activation of caspases (detectable in populations by enzymatic analysis or in single cells by staining) or the activation of endonucleases (detectable on population level as DNA-fragmentation). Moreover, a typical type of chromatin condensation (detectable by DNA stains) and the display of phosphatidylserine on the outside of the plasma membrane (detectable by Annexin staining) is highly correlated with apoptotic death. • **Advantages:** Adds mechanistic information to cytotoxicity data. Several endpoints are easy to quantify and useful for high through-put measurements. • **Disadvantages:** Not all types of cell death may be detected by a given endpoint. Needs to be combined with a general cytotoxicity test. Some endpoints are prone to artefacts (Annexin staining) and some staining techniques (TUNEL, caspase-3) lead to an un-intentional selection of subpopulations. Caspase activity measurement does not easily yield a prediction model for the extent of cell death.

3. Cell growth	
Cell counting	For some cell populations, continued growth is a defining feature, and thus impaired growth needs to be considered as a reduction of viability. Notably, impaired growth/proliferation is not necessarily correlated with cell death; it is thus rather a functional viability endpoint than a cytotoxicity measure. A special case for growth is the increase in cell size without proliferation. This feature is e.g., seen for the extension of neurites by neurons. The gold standard analytical endpoint for the growth/proliferation endpoint is counting (or morphometry). There are many ways of counting cells, either as single particles (e.g., by FACS or HCI) or by assessing a biochemical parameter correlated to cell number (e.g., DNA content). • **Advantages:** Growth can be a sensitive parameter of cell well-being. • **Disadvantages:** Growth is a functional endpoint, not necessarily linked to cytotoxicity; artefacts for growth endpoints may arise from inhomogeneous growth of subpopulations: moreover, growth may hide ongoing cell death, and thus needs careful control in combination with cytotoxicity assays.
BrdU or EdU incorporation	Measures new DNA synthesis based on incorporation of the easily detectable nucleoside analogs BrdU (or EdU) into DNA. BrdU can be detected e.g., by fluorescent-labelled antibodies in permeabilised cells. Alternatively, radiolabelled thymidine can be used. • **Advantages**: Measurement on single cell level. Easy to quantify and use at high throughput. • **Disadvantages**: BrdU/EdU can be cytotoxic; no information available on how often one given cell has divided. High cost and effort compared to counting.
Staining of cellular components that are proportional to overall cell mass (proteins by e.g., sulforhodamine B or crystal violet; DNA by Hoechst H-33342)	These assays evaluate a surrogate measure of overall cell mass and assume that it correlates with total cell number. In non-proliferating cells, or with continuous ongoing proliferation, the endpoints are also frequently used as indicators of cytotoxicity, as dead cells often detach from plates and reduce the overall cell mass. • **Advantages**: Simple and cheap; lots of historical data. • **Disadvantages:** Mostly not a single cell measure but only population level. Protein staining is only a surrogate endpoint of real cell number. For DNA quantification with Hoechst 33342: fluorescent probe penetration, bleaching, and cytotoxicity are issues to be considered. Crystal violet is a toxic/CMR classified chemicals and the use should be restricted.

4. Cellular metabolism	
3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium (MTT) assay, or similar tetrazolium dye reduction assays (e.g., WST-8)	Biochemical activity (mostly mitochondrial metabolism; production of reducing equivalents like NAD(P)H) in viable cells causes reduction of the tetrazolium dye. The resultant formazan is extracted and measured spectrophotometrically. The rate of formation of formazan corresponds to the function of essential cellular processes like respiration. • **Advantages:** High throughput, easy, robust, low cost. Used in several ISO standards and OECD test guidelines. High sensitivity. Can be used for tissue constructs. • **Disadvantages:** Measures amount of viable cells (only indirect measure of cell death), and needs control for contribution of proliferation. Cells with reduced mitochondrial function may appear non-viable. Inhibition of cell metabolism by the test item causes low values in the assay which is not necessarily related to cell viability. Some test items interfere with the assay e.g., by reducing the dye why interference testing is recommended. Measurement usually not on single cell level. Some cell cultures need long time to reduce sufficient amount of dye (no sharp time point for viability definition). Assessment of kinetic of the reduction may be necessary to ensure proper selection of incubation time with a tetrazolium dye (to avoid reaching plateau of OD).
Resazurin reduction assay (sometimes called Alamar blue)	Similar to tetrazolium reduction assays. Fluorescent/absorbent resorufin is formed from resazurin through mitochondrial metabolism of viable cells. • **Advantages:** Many tests can be performed rapidly in multi-well dishes. Cells can be tested repeatedly (non-invasive measurement). High sensitivity. • **Disadvantages:** Cells with reduced mitochondrial function may appear non-viable. Some test items interfere with the assay (e.g., superoxide also reduces the dye) why interference testing is recommended. Measurement only on population level. Some cell cultures need a long time to reduce sufficient resazurin (no sharp time point for viability definition).
Mitochondrial depolarisation assays (based on fluorescent indicator dyes)	Many organelle functions are used as endpoints of general cell health. Most frequently, mitochondrial function is assessed (see MTT, resazurin). One mitochondrial test on the single cell level is the measurement of mitochondrial membrane potential by addition of potential sensing fluorescent dyes like JC-1, TMRE, MitoTracker, etc. Quantification is by HCI or FACS • **Advantages**: Fast, inexpensive, high throughput; single cell information.

	• **Disadvantages**: As for MTT (measures cell function, not cytotoxicity). Artefacts by test items that affect mitochondria specifically. Artefacts by test items that affect plasma membrane potential. Artefacts due to bleaching, quenching and unquenching, and due to shape changes and clustering of mitochondria.
Neutral Red Uptake (NRU) (ISO 10993)	A cell organelle function assay assessing lysosomal function. Active cells accumulate the red dye in lysosomes and the dye incorporation is measured by spectrophotometric analysis. • **Advantages:** Low cost. Used in several ISO standards and OECD test guidelines. Historic data base. • **Disadvantages:** Normalisation required for quantitative measurement, e.g., with protein content or number of cells. Usually only gives information at the population level. Not suited for tissue constructs and certain cell lines. Not suitable for test items that affect lysosome function.
ATP assays	Measurement of the total ATP content in a cell population. Dying cells fail to produce ATP, have an increased ATP consumption, and may lose ATP through perforations of the plasma membrane. For the test, cell lysates are prepared, and the ATP content is assesses by a luminometric assay. • **Advantages:** Fast, high throughput • **Disadvantages:** No single cell data, expensive, requires a luminometer, as MTT: measurement of viable cell mass, not a direct measure of cytotoxicity. Artefacts as for other mitochondrial tests.

Glossary

All terms and their descriptions should be considered as working definitions for the purpose of this Guidance Document only.

Acceptance criteria: Criteria for when results can be accepted, i.e. a set of well-defined parameters describing aspects of the *in vitro* method such as range for positive and negative controls.

Accuracy: Refers to the closeness of a measured value to a standard or known value.

Authentication: Authentication of a cell line is the sum of the process by which a line's identity is verified and shown to be free of cross-contamination by other cell lines and/or contamination caused by bacteria, yeast or fungi, mycoplasm.

Adverse Outcome Pathway (AOP): An analytical construct that describes a sequential chain of causally linked events at different levels of biological organisation that lead to an adverse human health or environmental effect.

American National Standards Institute (ANSI): A private non-profit organisation that oversees the development of voluntary consensus standards for products, services, processes, systems, and personnel in the United States. https://www.ansi.org/

American Society for Testing and Materials International (ASTM): An international standards organisation that develops and publishes voluntary consensus technical standards for a wide range of materials, products, systems, and services. https://www.astm.org/

Amplicon: A piece of DNA or RNA that is the source and/or product of natural or artificial amplification or replication events. It can be formed using various methods including Polymerase Chain Reactions (PCR), ligase chain reactions, or natural gene duplication.

Apoptosis: Process of programmed cell death generally characterised by distinct morphological characteristics and energy-dependent biochemical mechanisms. Apoptosis is considered an essential component of various processes including normal cell turnover, proper development and functioning of the immune system, hormone-dependent atrophy, embryonic development and chemical-induced cell death.

Archive: A designated area or facility (e.g., cabinet, room, building or computerised system) for the secure storage and retention of records and materials.

Assay: A defined laboratory procedure for qualitatively or quantitatively measuring the presence or amount or the functional activity of a target or analyte. An assay can be considered as a technical operation that consists of determination of one or more characteristics of a given product, process or service according to a specified procedure.

Batch/Lot: A specific quantity of a test item, reference item or test system such as cells, tissues, assay reagent or other consumable, produced during a defined cycle in such a way that it could be expected to be of a uniform character and should be designated as such.

BenchMark Dose (BMD) or Concentration (BMC): The dose or concentration associated with a pre-specified biological response. It was developed as an alternative to the use of No Observed Adverse Effect Level (NOAEL) and Lowest Observed Adverse Effect Level (LOAEL).

Best practice: Generally accepted optimal methods or techniques that have consistently shown superior results among different labs as compared to those achieved with other means, and that is used as a benchmark. The term is also used to describe the process of developing and following a standard way of doing things that multiple organisations can use. Best practices are a snapshot and can be subject to change based on ongoing scientific dialogue and advancements.

Between-laboratory assessment: Phase of method validation, also often referred to as between (or inter) laboratory validation, in which different operators in different laboratories perform (or run) the *in vitro* method independently to establish whether or not an *in vitro* method can be successfully established in different laboratories, e.g., to assess the between laboratory reproducibility (BLR).

Biokinetics: Time-course of a chemical (substance and mixture) and its metabolites in a living organism, i.e., increase or decrease of substance concentration at the site of measurement due to transport or due to formation or breakdown.

Biological pathway: A series of actions among molecules in a cell that leads to a certain product or a change in the cell. Such a pathway can trigger the assembly of new molecules, such as a lipid or protein. Pathways can also turn genes on and off, or spur a cell to move. Some typical types of biological pathways are metabolic pathways and signalling pathways.

Carcinogenicity: The property of any agent (chemical, physical or biological agent) directly involved in causing cancer (carcinogen). Carcinogenicity results in an increased incidence of tumours, increased proportion of malignant tumours or a reduction in the time to appearance of tumours, compared with concurrent control groups. The process of carcinogenesis involves the transition of normal cells into cancer cells via a sequence of stages that may entail both genetic alterations (i.e. mutations) and non-genetic events.

Coefficient of Variation (CV): A measure of spread that describes the amount of variability relative to the mean. Because the coefficient of variation is per definition unrelated to the magnitude of the mean and also unitless, it can be used instead of the standard deviation to compare the spread of data sets that have different units or different means.

Comparative Genomic Hybridisation analysis (aCGH): A molecular cytogenetic method for analysing copy number variations relative to ploidy level in the DNA of a test sample compared to a reference sample, without the need for culturing cells. The aim of this technique is to quickly and efficiently compare two genomic DNA samples arising from two sources, which are most often closely related, because it is suspected that they contain differences in terms of either gains or losses of either whole chromosomes or subchromosomal regions (a portion of a whole chromosome).

Computerised system: A computerized system is a function (process or operation) integrated with a computer system and performed by trained personnel. The function is

controlled by the computer system. The controlling computer system is comprised of hardware and software. The controlled function is comprised of equipment to be controlled and operating procedures performed by personnel.

Cytotoxicity: General cytotoxicity (or basal cytotoxicity) is the result of toxic effects on structures and functions common to all cells of the body, such as DNA, chromosomes, mitochondria, the cytoskeleton and various membranes.

Data (derived data): Derived data depend on raw data and can be reconstructed from raw data (e.g., final concentrations as calculated by a spreadsheet relying on raw data, result tables as summarised by a LIMS, etc.).

Data (raw data): Data (raw data) may be defined as measurable or descriptive attribute of a physical entity, process or event. The GLP Principles define raw data as all laboratory records and documentation, including data directly entered into a computer through an automatic instrument interface, which are the results of primary observations and activities in a study and which are necessary for the reconstruction and evaluation of the report of that study.

Data integrity: The extent to which all data are complete, consistent and accurate throughout the data lifecycle.

Data lifecycle: All phases in the life of the data (including raw data) from initial generation and recording through processing (including transformation or migration), use, data retention, archive / retrieval and destruction (if applicable).

Defined Approach to Testing and Assessment: A defined approach consists of a fixed data interpretation procedure (DIP) (e.g. statistical, mathematical models) applied to data (e.g., in silico predictions, in chemico, in vitro data) generated with a defined set of information sources to derive a prediction. In contrast to the assessment process within Integrated Approaches to Testing and Assessment (IATA), that necessarily involves some degree of expert judgment, predictions generated with defined approaches are rule-based and can either be used on their own if they are deemed to fit-for-purpose or considered together with other sources of information in the context of IATA.

Design Qualification (DQ): Documented verification that the proposed design of the facilities, equipment, or systems is suitable for the intended purpose. This definition is applicable for complex instrumentation (computerised systems).

Effective Concentration 50 (EC_{50}) and Inhibition Concentration 50 (IC_{50}): In *in vitro* cell and tissue culture, EC_{50} is the concentration causing a half-maximal response for any measured biological effect of interest, and is equivalent to median effective dose (ED_{50}) and median lethal dose (LD_{50}) used in animal experiments. EC_{50} is used for read-outs that increase with concentration, whereas IC_{50} is used in case of an *in vitro* method where there is a decline in read-out. IC_{50} is therefore the test item concentration causing a reduction of response/binding etc. by half.

Emulsion: A stable dispersion of liquid droplets in another liquid, where the two are immiscible.

Engelbreth-Holm-Swarm (EHS) gel: Gelatinous protein mixture secreted by an EHS mouse sarcoma cells which are rich source of basement membrane components often used in cell and tissue culture work.

EURL ECVAM Database service on Alternative Methods (DB-ALM): A database providing comprehensive descriptions of non-animal methods together with related

information. Method descriptions are provided at two levels of detail, such as Summary Descriptions in an OECD compliant format (OECD, 2014) and/or detailed Standard Operating Procedures. DB-ALM originates from a requirement for EURL ECVAM to establish and manage public databases on alternative approaches as described in Annex VII of Directive 2010/63/EU on the protection of animals used for scientific purposes. https://ecvam-dbalm.jrc.ec.europa.eu

EURL ECVAM Scientific Advisory Committee (ESAC): Advises the European Union Reference Laboratory (EURL) ECVAM on scientific issues. ESAC's main role is to conduct independent peer reviews of validation studies of alternative test methods, assessing their scientific validity for a given purpose.

European Union Reference Laboratory on Alternatives to Animal Testing (EURL ECVAM): ECVAM was established in 1991 pursuant to a requirement in Directive 86/609/EEC that the European Commission (EC) and its member states actively support the development, validation, and acceptance of methods to replace, reduce, or refine the use of animals in laboratories. The activities of ECVAM were assumed by the European Union Reference Laboratory on Alternatives to Animal Testing (EURL ECVAM), which was formally established in 2011 as the Union Reference Laboratory specified in section 47, Article 48, and Annex VII of the European Commission's Directive 2010/63/EU. https://eurl-ecvam.jrc.ec.europa.eu/

European Chemicals Agency (ECHA): Agency of the European Union (EU) that manages technical, scientific and administrative aspects of EU chemicals legislation, notably the regulation on the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH). https://echa.europa.eu/

European Directorate for the Quality of Medicines & HealthCare (EDQM): Organisation that is responsible for the European Pharmacopoeia and the European Biological Standardisation Programme. https://www.edqm.eu/

European Food Safety Authority (EFSA): Agency of the European Union that provides independent scientific advice in the fields of food and feed safety, animal health and welfare, plant protection and plant health and communicates on existing and emerging risks associated with the food chain. http://www.efsa.europa.eu/

European Medicines Agency (EMA): Agency of the European Union that is responsible for the protection of public and animal health through the scientific evaluation and supervision of medicines. http://www.ema.europa.eu/

European Union Network of Laboratories for the Validation of Alternative Methods (EU-NETVAL): A network of highly qualified laboratories to (1) respond to some of the provisions of Directive 2010/63/EU, (2) generate *in vitro* method information that is reliable, relevant and based on current best quality and scientific practices, (3) increase the European Commission's validation capacity of *in vitro* methods and (4) provide a laboratory network knowledgeable on the routine implementation of good *in vitro* method practices for regulatory use in human safety assessment. https://eurl-ecvam.jrc.ec.europa.eu/eu-netval

Foetal Bovine/Calf Serum (FBS/FCS): Foetal bovine serum or often referred to as foetal calf serum is the liquid fraction of clotted blood (depleted of cells, fibrin and clotting factors, but containing a large number of nutritional (e.g., amino acids, sugars, lipids) and macromolecular factors (e.g., growth factors and hormones) considered by a large community to be essential for cultured cell growth. FBS/FCS is derived from the

blood drawn from a bovine foetus via a closed system of collection at the slaughterhouse. It is the most widely used growth supplement for cell and tissue culture media because of its high content of embryonic growth promoting factors and its low level of antibodies. When used at appropriate concentrations it may supply many defined but also undefined components that have been shown to satisfy specific metabolic requirements for the culture of cells and tissues.

Genetically Modified Organisms (GMOs): An organism in which the genetic material has been altered in a way that does not occur naturally by mating and/or natural recombination.

Good Cell Culture Practice (GCCP): Guidelines developed in 2005 to define minimum standards in cell and tissue culture work. This GCCP guidance lists a set of six principles intended to support best practice in all aspects of the use of cells and tissues *in vitro*, and to complement, but not to replace, any existing guidance, guidelines or regulations.

Good Laboratory Practice (GLP): A quality system concerned with the organisational process and the conditions under which non-clinical health and environmental safety studies are planned, performed, monitored, recorded, archived and reported. GLP in this document refers to the OECD Principles of GLP (OECD, 1998a) unless otherwise stated.

Hazard: An intrinsic feature of a stressor (e.g., chemical or physical in nature) to cause harm or adverse effects to human health and to the environment. It is a qualitative (for example in the case of classifications) or quantitative expression of the adverse effects elicited by a test item under defined conditions of exposure.

High Performance Liquid Chromatography (HPLC): High performance liquid chromatography (or high-pressure liquid chromatography) is a chromatographic technique that can separate a mixture of compounds when in solution and is used in biochemistry and analytical chemistry to identify, quantify and purify the individual components of the mixture.

High-Efficiency Particulate Air (HEPA): A type of air filter, also sometimes called high-efficiency particulate arrestance, with a minimum efficiency rating of 99.97% for the removal of 0.3 µm diameter or larger particulate matter[1].

High-Throughput Screening (HTS): A scientific approach relevant to chemistry and biology in which a very large number (e.g., tens of thousands per day) of experimental samples are subjected to testing under given conditions in a prescribed procedure.

***In silico*:** The technique of performing experiments via computer simulations. Examples are Structure-Activity Relationships (SAR) and Quantitative Structure-Activity Relationships (QSAR).

***In vitro*:** The technique of performing a given experiment in a test tube, or, more generally, in a controlled environment outside of a living organism.

***In Vitro* to *In Vivo* Extrapolation (IVIVE):** The qualitative or quantitative transposition of experimental results or observations made *in vitro* to predict phenomena *in vivo*, i.e. in whole organisms.

***In vivo* :** Experimentation using a whole, living organism as opposed to a partial or dead organism, or an *in vitro* controlled environment. Animal testing and clinical trials are two forms of *in vivo* research.

Inhibitor or spiked up control: Mix of test item and positive control to assess any effect of inhibition of the test item on the test system endpoint measurements.

Installation Qualification (IQ): The documented verification that the facilities, systems and equipment, as installed or modified, comply with the approved design and the manufacturer's recommendations. This definition is applicable for complex instrumentation (computerised systems).

Integrated Approaches to Testing and Assessment (IATA): IATA are pragmatic, science-based approaches for chemical hazard characterisation that rely on an integrated analysis of existing information coupled with the generation of new information using testing strategies.

Integrated Testing Strategies (ITS): ITS provide guidance on how various types of available data (including those obtained from *in vitro* testing methods or assays) should be evaluated, and addresses additional aspects on some elements such as the use of other toxicity data or weight of evidence analysis of existing and relevant data.

Intellectual Property Rights (IPR): The rights given to persons over the creations of their minds. They usually give the creator an exclusive right over the use of his/her creation for a certain period of time (WTO)[2].

Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM): ICCVAM is a permanent committee of the NIEHS under the National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods (NICEATM). ICCVAM is composed of representatives from 16 U.S. Federal regulatory and research agencies that require, use, generate, or disseminate toxicological and safety testing information.https://ntp.niehs.nih.gov/pubhealth/evalatm/iccvam/index.html

International Cell Line Authentication Committee (ICLAC): ICLAC is a voluntary, independent scientific committee that aims to make cell line misidentification more visible and promote authentication testing to combat this problem. http://iclac.org/

International Council for Harmonisation (ICH): ICH's mission is to achieve greater harmonisation worldwide to ensure that safe, effective, and high quality medicines are developed and registered in the most resource-efficient manner. The ICH brings together the receiving authorities and pharmaceutical industry to discuss scientific and technical aspects of drug registration. Harmonisation is achieved through the development of ICH Guidelines via a process of scientific consensus with regulatory and industry experts working side-by-side. Key to the success of this process is the commitment of the ICH regulators to implement the final Guidelines. http://www.ich.org/home.html

International Organization for Standardization (ISO): ISO is an independent, non-governmental international organisation that brings together experts to share knowledge and develop voluntary, consensus-based, market relevant International Standards that support innovation and provide solutions to global challenges. https://www.iso.org/

International Uniform ChemicaL Information Database (IUCLID): A software application designed to capture, store, maintain and exchange data on intrinsic and hazard properties of chemicals (substances and mixtures). The freely downloadable tool assists chemical companies globally to fulfil their obligation to submit data to the EU Chemicals Agency (ECHA) under the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) Regulation (Regulation (EC) No 1907/2006).

***In vitro* method developer**: The person or entity who develops an *in vitro* method destined for regulatory use in human or environmental safety assessment.

ISO 9000: The ISO 9000 family of quality management standards is designed to help organisations ensure that they meet the needs of customers and other stakeholders while meeting statutory and regulatory requirements to a product or program. ISO 9000 deals with the fundamentals of quality management systems.

Japanese Center for the Validation of Alternative Methods (JaCVAM): Promotes the 3Rs in animal experiments for the evaluation of chemical substance safety in Japan and establishes guidelines for new alternative experimental methods through international collaboration. http://www.jacvam.jp/en/

Korean Centre for the Validation of Alternative Methods (KoCVAM): KoCVAM was established in 2009 as part of the National Institute of Food and Drug Safety Evaluation (NIFDS) under the Korean Food & Drug Administration (KFDA). In March 2013, the KFDA was restructured and renamed as the Ministry of Food and Drug Safety (MFDS). http://www.nifds.go.kr/en/inter/kocvam.jsp

Lactate DeHydrogenase (LDH): An enzyme that helps the process of turning sugar into energy for cells to use. LDH is present in many kinds of organs and tissues throughout the body, including the liver, heart, pancreas, kidneys, skeletal muscles, brain, and blood cells. An LDH assay is a means of measuring either the number of cells via total cytoplasmic LDH or membrane integrity as a function of the amount of cytoplasmic LDH released into the medium.

Limit Of Detection (LOD): The LOD is the lowest quantity of a substance that can be distinguished from the absence of that substance (a blank value) within a stated confidence limit (generally 1%).

Limits Of Quantification (LOQ): The Lower Limit Of Quantification (LLOQ) and Upper Limit Of Quantification (ULOQ) are the lowest and highest quantity of a substance that can be quantitatively determined with a stated acceptable precision and accuracy, under stated experimental conditions.

Lipophilicity: The ability of a chemical (substance and mixture) to dissolve in non-polar environments such as oils, lipid membranes, and non-polar solvents such as hexane or toluene.

Mass Spectrometry (MS): An analytical technique that measures the mass-to-charge ratio of charged particles. It is used for determining masses of particles, for determining the elemental composition of a sample or molecule, and for elucidating the chemical structures of molecules such as peptides and other chemical compounds.

Medical device: An article, instrument, apparatus or machine that is used in the prevention, diagnosis or treatment of illness or disease, or for detecting, measuring, restoring, correcting or modifying the structure or function of the body for some health purpose. Typically, the purpose of a medical device is not achieved by pharmacological, immunological or metabolic means (WHO[3]).

Metadata: Metadata are data that describe the attributes of other data, and provide context and meaning. Typically, these are data that describe the structure, data elements, inter-relationships and other characteristics of data. They also permit data to be attributable to an individual.

Method endpoint: Quantitative or quantitative measurable characteristics that serve as indicators of a putatively pathologic process or related biochemical or molecular events, e.g., measured absorbance in a cytotoxicity assay or a skin irritation *in vitro* method.

Me-too test method: A colloquial expression for a test method that is structurally and functionally similar to a validated and accepted reference test method. Such a test method would be a candidate for catch-up validation (OECD, 2005).

Microorganism: Any microbiological entity, cellular or non-cellular, capable of replication or of transferring genetic material, including viruses, viroids, animal and plant cells in culture.

Minimal Essential Medium (MEM): A synthetic cell culture media for *in vitro* cell and tissue culture work, developed by Harry Eagle, one of the most widely used synthetic cell culture media.

Minimum Significant Ratio (MSR): Parameter that can be used to quantify assay reproducibility and resolution (the smallest ratio between two compound potencies which can be detected in the *in vitro* method).

Ministry of Agriculture, Forestry and Fisheries (MAFF): Japanese Ministry related to agricultural, forestry and fisheries products, covering from production to consumption and also to rural development and promotion of the welfare of rural inhabitants with a view to achieving a stable supply of food, sound development of the agriculture, forestry and fisheries industries and upgrading of the welfare of rural inhabitants. http://www.maff.go.jp/e/

Ministry of Health, Labour and Welfare (MHLW): Japanese Ministry responsible for the approval and administration of drugs, medical devices and cosmetics in Japan. http://www.mhlw.go.jp/english/

Mixture: A combination of two or more chemicals (liquid or solid) that do not react with each other.

Molecular Initiating Event (MIE): The initial interaction between a molecule and a biomolecule or biosystem that can be causally linked to an outcome via a pathway.

Multi-constituent substance: A substance, defined by its quantitative composition, in which more than one main constituent is present in a concentration ≥ 10% (w/w) and < 80% (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

Mutual Acceptance of Data (MAD): The OECD MAD is a multilateral agreement which states that test data generated in any member country in accordance with OECD Test Guidelines and GLP shall be accepted in other member countries for assessment purposes and other uses relating to the protection of human health and the environment. The application of MAD avoids unnecessary and costly duplication of testing as well as non-tariff barriers to trade. In addition, it reduces the number of laboratory animals used for *in vivo* testing.

Nanomaterial: A natural, incidental or manufactured material containing particles, in an unbound state or as an aggregate or as an agglomerate and where, for 50 % or more of the particles in the number size distribution, one or more external dimensions is in the size range 1 nm - 100 nm[4].

Negative control: Separate part of a test system treated with an item for which it is known that the test system should not respond; the negative control provides evidence that the test system is not responsive under the actual conditions of the assay.

Nephelometry: A technique for determining the amount of turbidity in a solution based upon the measurement of scattering of light.

OECD Harmonised Templates (OHTs): Standard data formats for reporting information used for the risk assessment of chemicals, mainly studies done on chemicals to determine their properties or effects on human health and the environment, but also for storing data on use and exposure.

OECD Test Guidelines (TG): OECD Test Guidelines are harmonised test methods included in the OECD Council Decision on Mutual Acceptance of Data. This means that "data generated in the testing of chemicals in an OECD Member country (or some non-member economies) in accordance with OECD Test Guidelines and OECD principles of Good Laboratory Practice shall be accepted in other Member countries (or non-member economies) for purposes of assessment and other uses relating to the protection of man and the environment".

Official Medicines Control Laboratories (OMCLs): European laboratory network supporting receiving authorities (including the issuing of guidelines) in the area of quality control of marketed medicinal products for human and veterinary use. https://www.edqm.eu/en/news/omcl-network

Omics: A general term for a broad discipline of science and engineering for analysing the interactions of biological information objects in various omes (these include genome, transcriptome, proteome, metabolome, expressome, and interactome).

Some examples of 'Omics' technologies:

- genomics
- proteomics
- metabolomics
- transcriptomics

Operational Qualification (OQ): Documented verification that all aspects of the facility, systems and equipment which can affect product quality; performs as intended throughout all anticipated operating ranges. This definition is applicable for complex instrumentation (computerised systems).

Particulates/Particulate Matter (PM): Tiny subdivisions of solid matter suspended in a gas or liquid (also known as particulate matter, fine particles and soot).

Performance Based Test Guideline (PBTG): A test guideline that contains one or more *in vitro* methods that are mechanistically and functionally similar. A PBTG defines the important components of the *in vitro* method and describes in detail characteristics and performance standards that a new *in vitro* method should meet in order to be considered as an additional method.

Performance Qualification (PQ): The documented verification that the facilities, systems and equipment, as connected together, can perform effectively and reproducibly, based on the approved process method and product specification. This definition is applicable for complex instrumentation (computerised systems).

Performance Standards (PS): The purpose of performance standards is to provide the basis by which new or modified *in vitro* methods, both proprietary (i.e. copyright, trademarked, registered) and non-proprietary, can be deemed to be structurally and mechanistically similar to a validated reference method and demonstrate to have sufficient reliability and relevance for specific purposes (i.e. in accordance with the principles to OECD GD 34).

Physiologically-Based ToxicoKinetic (PBTK[5]) models: Physiologically based toxicokinetic, or alternatively referred to as physiologically based pharmacokinetic or biokinetic models, are quantitative descriptions of absorption, distribution, metabolism, and excretion (ADME, possibly including toxicity as ADMET) of synthetic or natural chemical substances in humans and other animal species. PBTK models are increasingly being used as an effective tool for designing toxicology experiments and for conducting extrapolations essential for risk assessments (e.g., in pharmaceutical research and drug development, and in health risk assessment for cosmetics or general chemicals).

Polymerase Chain Reaction (PCR): A molecular biology technique used to make multiple copies of a segment of DNA. PCR is very precise and can be used to amplify, or copy, a specific DNA target from a mixture of DNA molecules[6]. It is based on the natural process of DNA replication.

Population Doubling Level (PDL): Refers to the total number of times the cells in the population have doubled since their primary isolation *in vitro*, and are usually an estimate rounded off to the nearest whole number.

Positive control: Separate part of the test system treated with an item for which it is known that the test system should respond. The positive control provides evidence that the test system is responsive under the actual conditions of the assay. The positive control is endpoint specific to the test system.

Prediction model: The method by which the *in vitro* endpoint value(s) is used to predict the *in vivo* equivalent activity (i.e., degree of toxicity).

Proficiency chemicals (substances): A set of chemicals recommended within OECD Test Guidelines to be used by laboratories to demonstrate their technical proficiency prior to the routine use of a test method falling within the adopted OECD test guideline. These chemicals represent either a subset of the reference chemicals included in the Performance Standards relating to the OECD TG, or chemicals used in the validation studies of the test method falling within the OECD TG. Selection criteria for these test chemicals include, to the extent possible, chemicals that: i) represent the range of responses to be predicted, ii) have high quality reference data available; iii) cover the method's dynamic range of responses; iv) were correctly predicted by the test method during its validation study; v) cover a wide and representative range of relevant physical states, chemical classes, organic functional groups and structures falling within the applicability domain of the in vitro method; vi) are commercially available, and vii) are not associated with prohibitive acquisition and/or disposal costs.

Provenance: Describes the origin and culture history of a cell line, including its transfers among laboratories and repositories, its manipulation (physicochemical or genetic), tests for and the detection and elimination of contamination by other cell lines and/or contamination caused by bacteria, yeast or fungi, mycoplasma, genotypic and phenotypic characteristics, and verification of its identity.

Quality Assurance (QA)*:* All the planned and systematic actions by which adherence to laboratory testing standards, requirements, and record keeping procedures, and the accuracy of data transfer, are assessed by individuals independent of those performing the testing.

Quality assurance programme: A defined system, including personnel, which is independent of study conduct and is designed to assure test facility management of compliance with the Principles of Good Laboratory Practice. (OECD, 1998a).

Quality Control (QC): Operational techniques and activities that are used to fulfil given requirements for quality.

Quality Management System (QMS): Can be expressed as the organisational structure, procedures, processes and resources needed to implement quality management. GLP specifically refers to a quality system of management controls for test facilities and organisations to try to ensure the uniformity, consistency, reliability, reproducibility, quality, and integrity of test item non-clinical safety tests. Of all QMS regimes, the ISO 9000 family of standards is probably the most widely implemented worldwide.

Reagent: A substance or mixture for use in cell culture media, chemical analysis or other reactions.

Reference item: An article used to provide a basis for comparison with the test item.

Relevance: The term "Relevance" describes whether a procedure is meaningful and useful for a particular purpose.

Reliability: The term "Reliability" describes whether a procedure can be performed reproducibly within and between laboratories and over time.

Replace, Reduce, Refine (3Rs): A term describing current internationally accepted strategies for minimising use and suffering of laboratory animals used in experimental research. The optimal solution is to Replace the test method requiring animal experiments with one or several *in vitro* methods; if this is not possible at least it might be possible to modify the methods in order to Reduce the number of animals being used in each study without compromising data quality; if this is also not possible it might at least be possible to Refine the test method so that experiments are conducted in a way minimising stress and other impact on the animals.

Robustness: The insensitivity of test results to departures from the specified test conditions when conducted in different laboratories or over a range of conditions under which the test method might normally be used. If a test is not robust, it will be difficult to use in a reproducible manner within and between laboratories.

Saturation concentration: The maximum dissolved concentration of a test chemical that can be achieved under the test conditions.

Sensitivity: A measure of *in vitro* method performance that describes the proportion of all evaluated test items that are classified as positive for a particular toxicological endpoint, which are predicted as positive by the actual *in vitro* method. The terms "sensitivity" may also refer to *in vivo* tests when e.g., compared to human data.

Service Level Agreement (SLA): A contract between a service provider (either internal or external) and the end user that defines the level of service expected from the service provider.

Short Tandem Repeat (STR): Short sequences of DNA, usually of length 2-6 base pairs and directly adjacent to each other that are repeated numerous times along a given loci. They are also known as microsatellites. STRs are used to compare specific loci on DNA from two or more samples.

Signal Windows (SW): A measure of separation between maximum and minimum controls in an assay that accounts for the amount of variability in the assay (Sittampalam *et al*., 2017).

Single Nucleotide Polymorphism analysis (aSNP): Single nucleotide polymorphism or SNP (pronounced snip) analysis is a technique to detect a DNA sequence variation occurring when a single nucleotide - A, T, C, or G - in the genome (or other shared sequence) differs between members of a species (or between paired chromosomes in an individual). For example, two sequenced DNA fragments from different individuals, AAGCCTA to AAGCTTA, contain a difference in a single nucleotide.

Solid Phase MicroExtraction (SPME): A technique for separating mixtures of compounds without the use of solvents. SPME uses a fibre coated with a polymer or sorbent extracting phase that extracts chemical compounds from liquid or gas phases.

Solubility limit in water: The maximum attainable concentration in water at thermodynamic equilibrium between the aqueous pure phase and the solid (or liquid or gaseous) pure phase.

Specificity: A measure of *in vitro* method performance that describes the proportion of all evaluated test items that are classified as negative for a particular toxicological endpoint, which are predicted as negative by the actual *in vitro* method. The terms "specificity" may also refer to *in vivo* tests when e.g., compared to human data.

Sponsor: Means an entity which commissions, supports and/or submits a non-clinical health or environmental safety study.

Standard Deviation (SD): The expected squared deviation from the mean.

Standard Operating Procedure (SOP): A documented procedure which describes how to perform testing methods or assays or activities normally not specified in detail in study plans or test guidelines.

Standard Project Submission Form (SPSF): OECD standard form which specifies the information generally required to submit a proposal for new or updated Test Guidelines or related documents to the Working Group of the National Coordinators of the Test Guidelines Programme (WNT)[7], including the project description and the actions planned toward the development of the Test Guideline, the project milestones, and deliverables.

Structure-Activity Relationships and Quantitative Structure-Activity Relationships (SAR/QSAR): Structure-activity relationships and quantitative structure-activity relationships based on the chemical structure of a compound, collectively referred to as (Q)SARs, are simplified mathematical representations of complex chemical-biological interactions that can be used to predict the physicochemical and biological properties of molecules.

Study plan: A document which defines the objectives and experimental design for the conduct of the study, including amendments (i.e., an intended change to the study plan after the study initiation date).

Suspension: A stable dispersion of solid particles in a liquid.

Test item: An article that is the subject of a study (e.g., chemical, substance, nanomaterial, medical device, biologicals etc.). Test chemical may be used interchangeably for chemical based test items.

Test system: Any biological, chemical or physical system or a combination thereof used in a study. *In vitro* test systems are mainly biological systems (e.g., cells or tissues), although some of the more recent developments in alternatives to conventional *in vivo* testing (e.g., gene arrays for toxicogenomics) may also exhibit some attributes of physical-chemical test systems. Test kits, including proprietary test kits, should also be considered as test systems.

Testing method: A testing method, also known as assay, is a process or procedure used to obtain information on the characteristic of a substance or agent. Specific testing methods generate information regarding the ability of a substance or agent to produce a specific biological effect under specified conditions.

Toxicological endpoint: A direct marker of progression to an adverse outcome - e.g., morphological or physiological changes, functional impairments, disease symptoms or death - used to describe an adverse health effect (or a probability of that adverse effect) resulting from exposure to a test item. The test system response to an exposure of a test item may be measured by a series of endpoints. The most sensitive endpoint (critical endpoint) is the one that occurs at the lowest exposure level and associated with an adverse response (committed step).

Training set: A set of test items used to develop the prediction model for an assay. The training set items should have strong reference data (i.e., values from a recognised regulatory assay and derived from multiple runs of the reference tests) against which the *in vitro* assay endpoint values can be compared.

Untreated control: Test system that receives no treatment (e.g., no test chemical or solvent) but is processed concurrently and in the same way as the test system receiving the test item.

US Department of Agriculture (USDA): USDA is the U.S. federal executive department responsible for developing and executing federal laws related to farming, agriculture, forestry, and food. It aims to meet the needs of farmers and ranchers, promote agricultural trade and production, work to assure food safety, protect natural resources, foster rural communities and end hunger in the United States and internationally. https://www.usda.gov/

US Environmental Protection Agency (EPA): The EPA is an agency of the Federal government of the United States which was created for the purpose of protecting human health and the environment. https://www.epa.gov/

US Food and Drug Administration (FDA): The FDA is a regulatory and research agency within the US Department of Health and Human Services that is responsible for "protecting the public health by assuring the safety, efficacy and security of human and veterinary drugs, biological products [for humans], medical devices, the nation's food supply [including dietary supplements], cosmetics, tobacco products and products that emit radiation". https://www.fda.gov

Validation: The process by which the reliability and relevance of a procedure are established for a specific purpose (OECD, 2005).

Validation set: A set of selected chemicals used to assess the predictive capacity of an *in vitro* method based on the performance of the endpoint values by the reference *in vitro* method results. Testing of the validation set is a principal part of *in vitro* method validation.

Vehicle/solvent control: Separate part of a test system to which the vehicle/solvent for the test item is added without the test item; the vehicle/solvent control provides evidence for a lack of influence of the chosen vehicle/solvent on the test system under the actual conditions of the *in vitro* method.

Veterinary International Conference on Harmonization (VICH): VICH is a trilateral (EU-Japan-USA) programme aimed at harmonising technical requirements for veterinary product registration. Its full title is the International Cooperation on Harmonisation of Technical Requirements for Registration of Veterinary Medicinal Products. VICH was officially launched in April 1996. http://www.vichsec.org/

Within-laboratory assessment: Phase (of method validation), also often referred to as within laboratory (or intra) laboratory validation, in which one or more operators from the same laboratory perform (or run) the *in vitro* method independently and at different times to establish whether or not an *in vitro* method meets established criteria e.g., to assess within-laboratory reproducibility (WLR).

World Health Organisation (WHO): The WHO was established in 1948 as a specialised agency of the United Nations. WHO is made up of 193 Member States, most of which are also UN Member States. WHO's mission is "the attainment by all peoples of the highest possible level of health". http://www.who.int/

Xenobiotic: A chemical foreign to the biological system, structurally distinct from endogenous compounds present within the biological system. A xenobiotic may also be directly pharmacologically, endocrinologically, or toxicologically active, or undergo metabolism in target organisms such as to become biologically active or inactive.

Z-factor: A measure of the separation between solvent control and test item signal which takes into account the dynamic range of the *in vitro* method and the data variation associated with the signal and control measurements. It is suitable for *in vitro* method quality assessment.

Notes

1. See: https://www3.epa.gov/ttncatc1/dir1/ff-hepa.pdf
2. See: https://www.wto.org/english/tratop_e/trips_e/intel1_e.htm
3. See: http://www.who.int/medical_devices/definitions/en/
4. See: http://ec.europa.eu/environment/chemicals/nanotech/faq/definition_en.htm
5. PBTK models are regarded as being synonymous to PBPK (physiologically-based pharmacokinetic), PBBK (physiologically-based biokinetic) and PBK (physiologically-based kinetic) models (Bessems *et al*., 2014).
6. See: https://www.nature.com/scitable/definition/polymerase-chain-reaction-pcr-110
7. See: http://www.oecd.org/env/ehs/testing/testguidelinesprogrammefaqs.htm

References

Bessems, J. et al. (2014), "PBTK modelling platforms and parameter estimation tools to enable animal-free risk assessment", *Regulatory Toxicology and Pharmacology*, Vol. 68/1, pp. 119-139, http://dx.doi.org/10.1016/j.yrtph.2013.11.008. [2]

OECD (2017), *Guidance Document for Describing Non-Guideline In Vitro Test Methods*, OECD Series on Testing and Assessment, No. 211, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264274730-en. [5]

OECD (2005), *OECD Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment*, OECD Publishing. Paris. [4]

OECD (1998), *OECD Principles on Good Laboratory Practice*, OECD Series on Principles of Good Laboratory Practice and Compliance Monitoring, No. 1, OECD Publishing, Paris, http://dx.doi.org/10.1787/9789264078536-en. [3]

Sittampalam, G. et al. (2004), *Assay Guidance Manual*, Bethesda (MD): Eli Lilly & Company and the National Center for Advancing Translational. [1]

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

The OECD is a unique forum where governments work together to address the economic, social and environmental challenges of globalisation. The OECD is also at the forefront of efforts to understand and to help governments respond to new developments and concerns, such as corporate governance, the information economy and the challenges of an ageing population. The Organisation provides a setting where governments can compare policy experiences, seek answers to common problems, identify good practice and work to co-ordinate domestic and international policies.

The OECD member countries are: Australia, Austria, Belgium, Canada, Chile, the Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea, Latvia, Lithuania, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, the Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. The European Union takes part in the work of the OECD.

OECD Publishing disseminates widely the results of the Organisation's statistics gathering and research on economic, social and environmental issues, as well as the conventions, guidelines and standards agreed by its members.

OECD Series on Testing and Assessment

Guidance Document on Good In Vitro Method Practices (GIVIMP)

In the past several decades, there has been a substantial increase in the availability of in vitro test methods for evaluating chemical safety in an international regulatory context. To foster confidence in in vitro alternatives to animal testing, the test methods and conditions under which data are generated must adhere to defined standards to ensure resulting data are rigorous and reproducible. Good In vitro Method Practices (GIVIMP) for the development and implementation of in vitro methods for regulatory use in human safety assessment aims to help reduce the uncertainties in cell and tissue-based in vitro method derived chemical safety predictions. GIVIMP provides guidance for test method developers and end users of resulting data on key elementes of in vitro methods. GIVIMP tackles ten important aspects related to in vitro work: (1) Roles and responsibilities, (2) Quality considerations, (3) Facilities (4) Apparatus, material and reagents, (5) Test systems, (6) Test and reference/control items, (7) Standard operating procedures (SOPs), (8) Performance of the method, (9) Reporting of results, (10) Storage and retention of records and materials.

Consult this publication on line at ***https://doi.org/10.1787/9789264304796-en***.

This work is published on the OECD iLibrary, which gathers all OECD books, periodicals and statistical databases. Visit ***www.oecd-ilibrary.org*** for more information.

OECD*publishing*
www.oecd.org/publishing

ISBN 978-92-64-30479-6
97 2018 41 1 E

9 789264 304796