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.

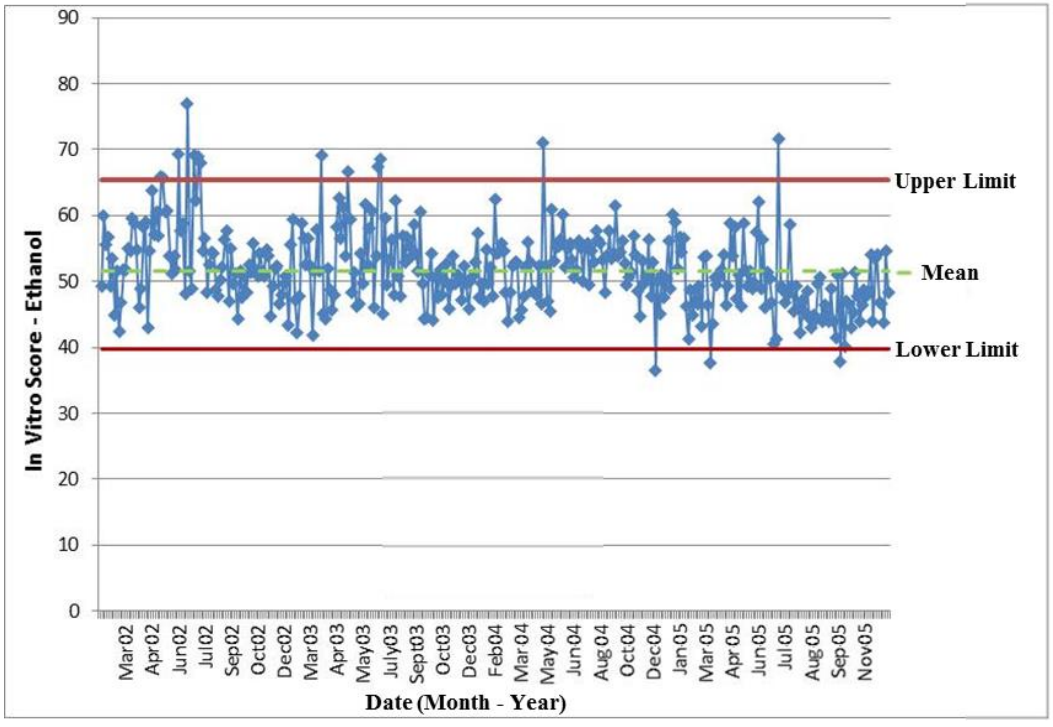
Monitoring and recording performance against negative and positive control items may constitute sufficient proof for the responsiveness of a given test system (OECD, 2004[1]). Non-response of the test system to the negative control and positive response to the positive control, within the acceptance 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.

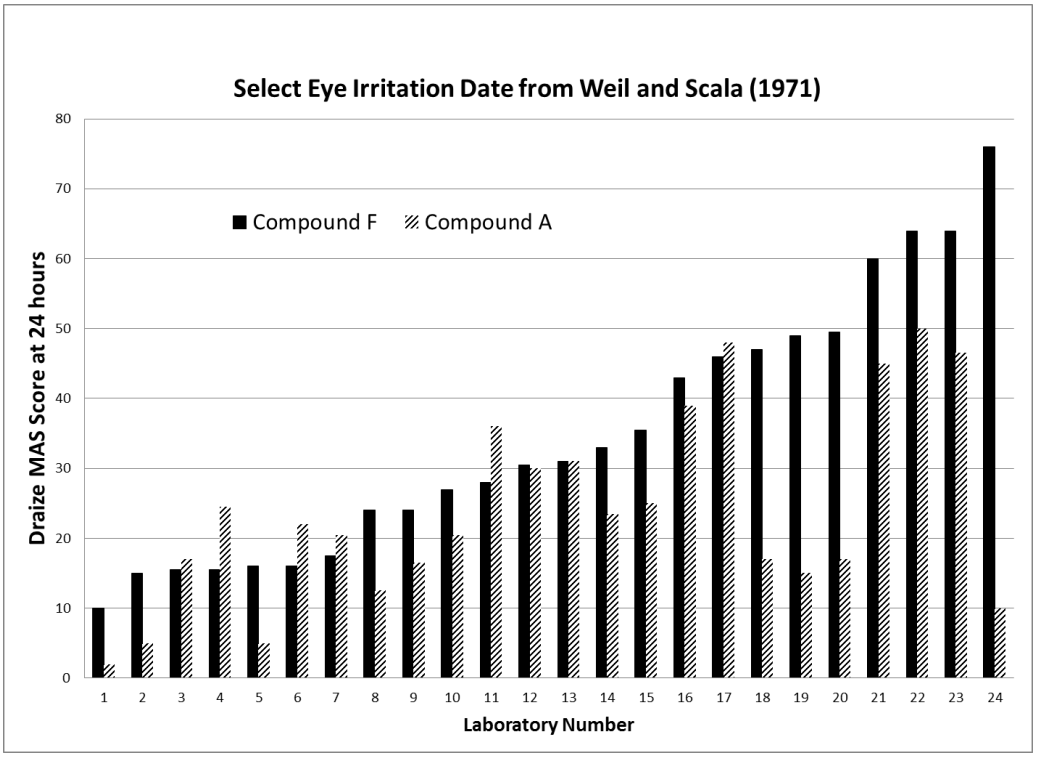
Figure 6.1. BCOP ethanol positive control QC chart March 2002 to November 2005



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.

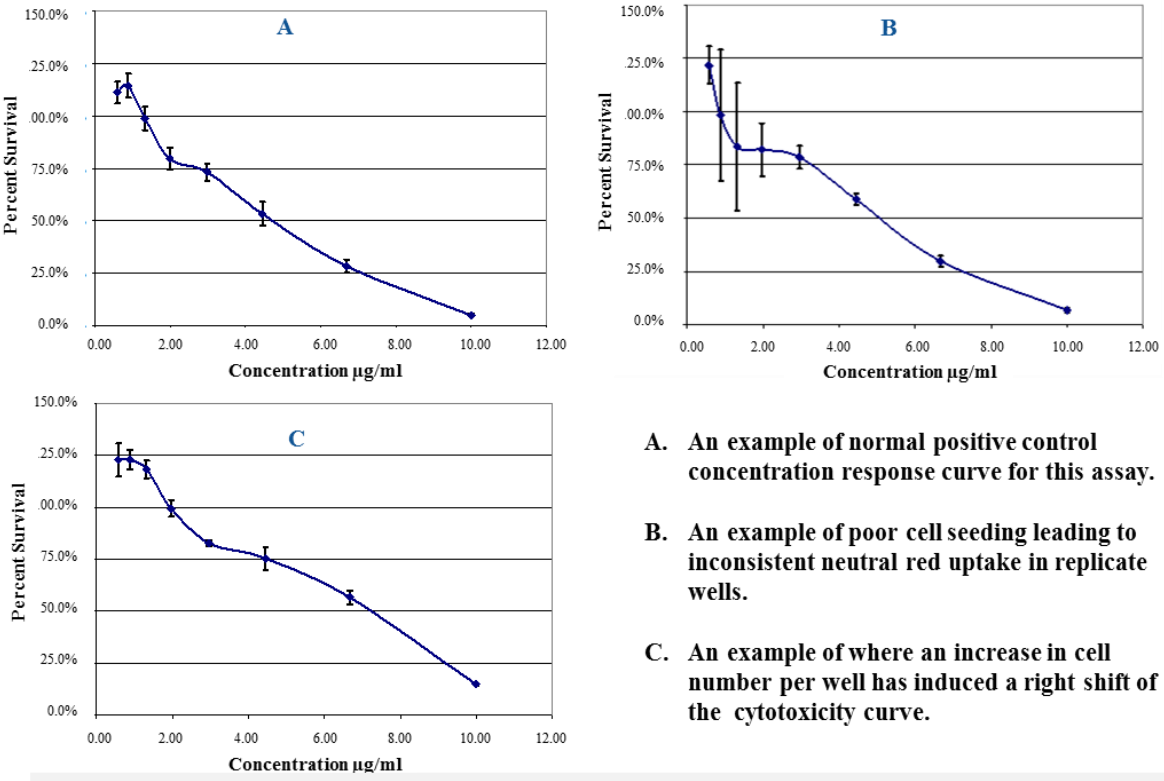
Figure 6.2. Historical data from the Draize Eye Irritation test on two chemicals at 24 hours after instillation where no positive control was included



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



A. An example of normal positive control concentration response curve for this assay.

B. An example of poor cell seeding leading to inconsistent neutral red uptake in replicate wells.

C. An example of where an increase in cell number per well has induced a right shift of the cytotoxicity curve.

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

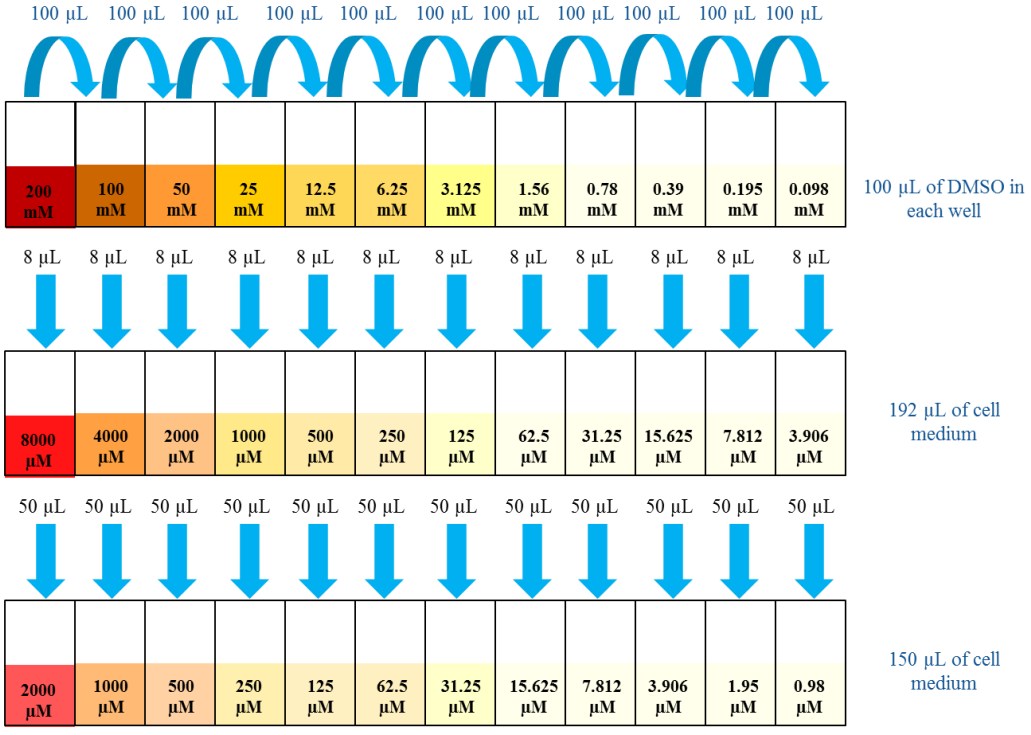
In many OECD TGs a preliminary test is carried out to determine the appropriate concentration (dose finding) range of the test item to be tested, and to ascertain whether the test item may have any solubility and cytotoxicity issues (e.g., TG 455, TG 442E).

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

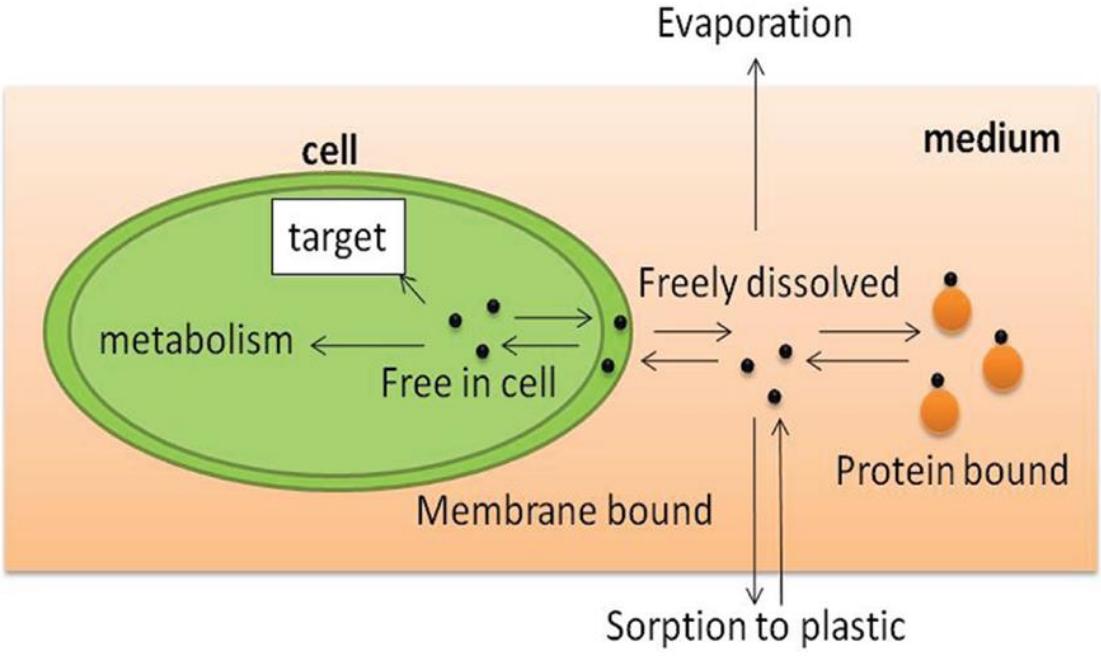
Just like the biokinetics *in vivo* are about what the body of the organism does to the test item, the biokinetics *in vitro* concerns what the *in vitro* test environment does to the test item. A central issue in biokinetics is that generally only the freely dissolved molecules of a chemical can pass 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]