Bench Marks!

Characteristics to Consider When Selecting a Positive Control Material for an *In Vitro* Assay

Elijah J. Petersen[1#], Andrew D. Nguyen[2#], Jeffrey Brown[2], John T. Elliott[1], Amy J. Clippinger[2], John Gordon[3], Nicole Kleinstreuer[4] and Matthias Roesslein[5]

[1]Biosystems and Biomaterials Division, Material Measurement Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA; [2]PETA Science Consortium International e.V., Stuttgart, Germany; [3]US Consumer Products Safety Commission (CPSC), Rockville, MD, USA; [4]National Institute of Environmental Health Sciences (NIEHS), National Toxicology Program Interagency Center for the Evaluation of Alternative Toxicological Methods, Research Triangle Park, NC, USA; [5]Empa, Swiss Federal Laboratories for Material Testing and Research, Particles-Biology Interactions Laboratory, St. Gallen, Switzerland

Abstract

The use of *in vitro* assays to inform decision-making requires robust and reproducible results across studies, laboratories, and time. Experiments using positive control materials are an integral component of an assay procedure to demonstrate the extent to which the measurement system is performing as expected. This paper reviews ten characteristics that should be considered when selecting a positive control material for an *in vitro* assay: 1) the biological mechanism of action, 2) ease of preparation, 3) chemical purity, 4) verifiable physical properties, 5) stability, 6) ability to generate responses spanning the dynamic range of the assay, 7) technical or biological interference, 8) commercial availability, 9) user toxicity, and 10) disposability. Examples and a case study of the monocyte activation test are provided to demonstrate the application of these characteristics for identification and selection of potential positive control materials. Because specific positive control materials are often written into testing standards for *in vitro* assays, selection of the positive control material based on these characteristics can aid in ensuring the long-term relevance and usability of these standards.

1 Introduction

A robust testing strategy is often built upon multiple sources of information, including existing data, physicochemical properties (e.g., pH, particle size, and distribution), non-testing sources (e.g., quantitative structure-activity relationships (QSAR) and read-across), and/or *in vitro* assays (Basketter et al., 2019). Importantly, mechanistic *in vitro* assays can query specific biological mechanisms or key events that lead to an adverse outcome and provide information about human-relevant effects (Avila et al., 2020). Positive control experiments are important to determine whether an *in vitro* assay is performing as expected and to gain confidence in the reliability and reproducibility of the test result.

Positive control experiments use positive control materials to generate a known, repeatable response of the endpoint being evaluated in the assay (Elliott et al., 2017; Petersen et al., 2020; Rösslein et al., 2015). For example, a cell cytotoxicity assay would have a positive control material that causes cell death

# contributed equally

Disclaimer: Certain commercial entities, equipment, or materials may be identified in this paper in order to adequately describe an experimental procedure or concept. Such identification is not intended to imply recommendation or endorsement by NIST, NIEHS or CPSC nor is it intended to imply that the commercial entities, equipment, or materials are necessarily the best available for the purpose. The findings and conclusions in this paper are those of the authors and do not necessarily represent the views of NIST, NIEHS or the CPSC.

Received February 11, 2021;
Epub February 24, 2021; © The Authors, 2021.

ALTEX 38(2), 365-376. doi:10.14573/altex.2102111

Correspondence: Elijah Petersen, PhD
Biosystems and Biomaterials Division, Material Measurement Laboratory
National Institute of Standards and Technology (NIST)
100 Bureau Drive, Gaithersburg, MD 20899, USA
(elijah.petersen@nist.gov)

(Elliott et al., 2017; Rösslein et al., 2015; Nelson et al., 2013; Leibrock et al., 2020), while an assay evaluating the potential for a chemical to cause inflammation would require a positive control material known to induce inflammation-like responses in a particular cell type (Petersen et al., 2020). Results from the positive control experiment can be used to develop specifications and provide a quantitative strategy to ensure confidence in the assay measurement system. The fundamental component of positive control experiments is a positive control material or set of positive control materials that provide the expected response in the test system.

Positive control materials should be fit-for-purpose in a measurement system. Depending on the assay, different characteristics of a positive control material may be more or less important. For example, both physicochemical properties (e.g., solubility in test media and stability in storage media) and pharmacological properties (e.g., estrogenicity and cytotoxicity) may need to be considered to identify an appropriate positive control material.

The use of positive control materials in *in vitro* assays is widely accepted, and recommendations on the selection of a positive control material have been discussed previously (OECD, 2018; Hartung et al., 2002). However, an in-depth rationale for the importance of certain characteristics and the problems that may occur if a positive control material does not exhibit these characteristics is often not provided. There are characteristics of a positive control material that may limit the information it provides, the transferability of an assay to other laboratories, or the ability to reproduce the results of an assay across time.

This discussion focuses on choosing a positive control material during assay development based on ten characteristics that can influence an assay's response to, and the long-term usability of, a positive control material: 1) the biological mechanism of action, 2) ease of preparation, 3) chemical purity, 4) verifiable physical properties, 5) stability, 6) ability to generate responses spanning the dynamic range of the assay, 7) technical or biological interference, 8) commercial availability, 9) user toxicity, and 10) disposability. Many of these characteristics should also be considered for proficiency test chemicals to demonstrate laboratory competency and increase the reproducibility of responses among laboratories (OECD, 2018). Assay developers often have to prioritize some characteristics over others when choosing among different potential positive control materials. Therefore, we present examples to highlight a decision process for selecting positive control materials and provide a case study evaluating the use of lipoteichoic acid (LTA) as a positive control material in the monocyte activation test (MAT).

2 Definitions

The following definitions are used to maintain consistency within this manuscript. These definitions are not intended to be recommendations for broader adoption.

- Fit-for-purpose: The positive control material is suitable or appropriate for testing the measurement system based on material characteristics such as physicochemical properties (e.g., stability or ease of handling) and biological activity.
- Dose-response function: The change in assay readout over varying concentrations of the test substance.
- Positive control experiment: An experimental procedure that uses a positive control material to reproducibly yield a measurable response that indicates a test method is performing as expected.
- Positive control material: A well-characterized reference material that can reproducibly yield a known response in a test method to indicate that the method is performing as expected (adapted from ISO, 2012).
- Reference material: Material, sufficiently homogeneous and stable with respect to one or more specified properties, which has been established to be fit for its intended use in a measurement process (ISO, 2015).
- Repeatability: The agreement among test results obtained within a single laboratory (intra-laboratory) when the procedure is performed on the same substance under identical conditions (OECD, 2005).
- Reproducibility: The agreement among results obtained from testing the same substance using the same test protocol between laboratories (inter-laboratory; adapted from OECD, 2005).
- Robustness: Resistance to measured change in the assay resulting from experimentally established or possibly unintended variations in experimental reagents, laboratory conditions, or protocols (adapted from Plant et al., 2014).
- Stability: The ability of a material to maintain a specific characteristic (e.g., biological response or chemical structure) for a specified period of time (adapted from ISO, 2012).

3 Relationship between reference materials and positive control materials

Reference materials are defined as materials that are stable and homogeneous with respect to specific properties that are relevant to a measurement process. These materials may play a key role in *in vitro* test methods by confirming that some aspect of the method is working as expected, such as verifying the performance of a plate reader. In the case of positive control materials, the property that should be stable and homogeneous is the response that the positive control material yields when evaluated in the assay, which often provides the basis for comparison when evaluating the response of other test materials. Once suitable positive control materials are identified, they are used routinely in positive control experiments to monitor the performance of the measurement system.

4 Considerations for selecting a positive control material

Important characteristics of a positive control material to consider are shown in Figure 1. These characteristics can be aggregated into three main themes: consistent performance (numbers 1 through 7), accessibility (number 8), and safety (numbers 9 and 10). If there is significant intra- or inter-laboratory variability in the positive control experimental results for a positive control material, one

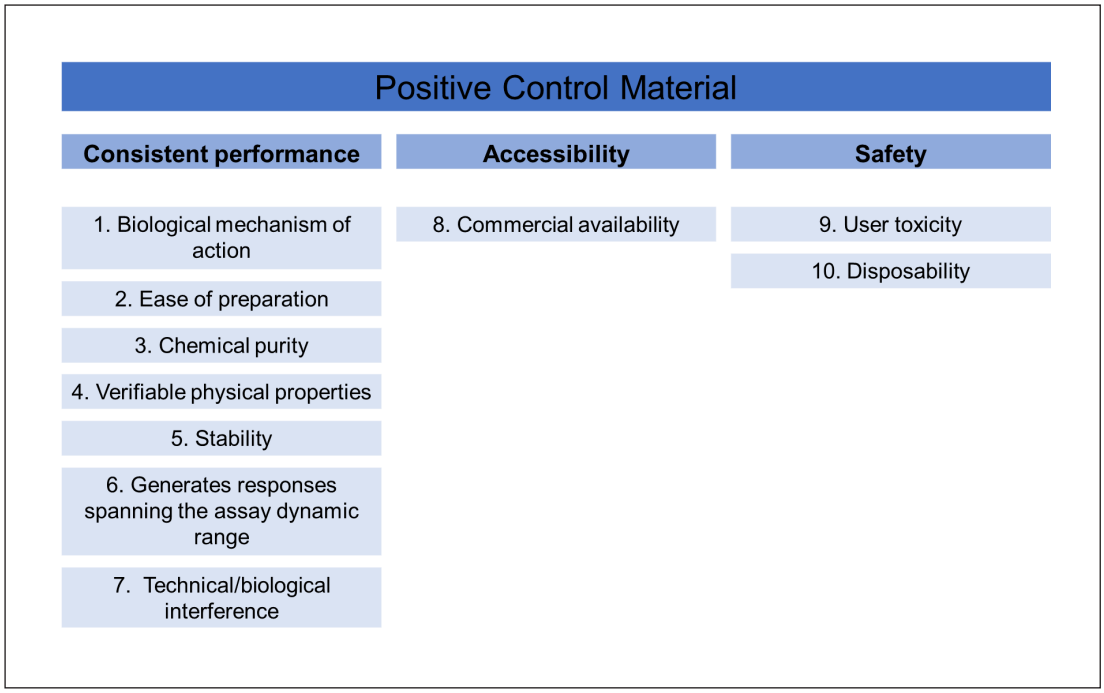

Fig. 1: Characteristics to consider when selecting a positive control material

of the characteristics in the consistent performance theme may be driving that variability. When choosing among possible positive control materials, fulfilling more characteristics within each theme suggests that the material is expected to achieve the objective of that theme more thoroughly than other materials. Overall, fulfilling characteristics in each of these themes will support the long-term usability and performance of the positive control material.

4.1 Biological mechanism of action

The positive control material should result in an assay response by activating the biological mechanism of action being monitored in the assay. Mechanisms can range from broad cellular processes to fine scale intermolecular interactions, and the level of complexity the assay is interrogating must be clear. For example, an assay may broadly assess cytotoxicity or it may be specific to apoptosis or necrosis (Petersen et al., 2020). For skin sensitization, covalent protein binding is a key event measured by a variety of assays (e.g., those in OECD, 2019), but can be broken down into several mechanisms, such as Michael acceptors, acetylating agents, SN2 agents, and Schiff bases (Chipinda et al., 2010; Urbisch et al., 2015).

Multiple positive control materials can be used in *in vitro* assays to parse out broad mechanistic endpoints as well as discriminate between finer mechanisms. When multiple positive control materials are used, it is possible to test multiple specific mechanisms and a large dynamic range of positive responses (i.e., strong, medium, and weak responses from different positive control materials). For example, well studied pathways, such as estrogenic and androgenic modulation, provide the opportunity to select from a range of potential positive control compounds to provide reference points along the biological response continuum for test material comparison (Browne et al., 2015; Kleinstreuer et al., 2017). However, a single positive control material relevant to the mode of action of the test chemical is commonly used, and the assumption is made that results for that material will be sufficient to determine if the assay results will be comparable across experiments and laboratories.

Another example of when the use of multiple positive control materials may be advantageous is for the testing of engineered nanomaterials (Barosova et al., 2020). In one study (Barosova et al., 2020), following exposure to a nanomaterial known to induce toxicity, the expected induction of inflammatory markers was not observed in the test system. Therefore, a second chemical (non-nanoparticle) positive control material was included in the experiment to determine whether the issue was the nanomaterial or the test system itself. The positive control chemical also did not produce the expected inflammatory response, ruling out the possibility of some nano-specific issue (e.g., deposition efficiency) and showing a more general issue with the system. This information led to identification of a proprietary ingredient in the medium used for culturing the tissue model that was inhibiting inflammation and, following its removal, the nanomaterial and chemical positive control materials showed the expected inflammatory response. This example highlights the importance of using a positive control material that acts through the same mechanism of action as the test substances (in this case the induction of inflammation) and the potential need to include an additional positive control material to rule out indirect complications (e.g., deposition efficiency of nanomaterials).

4.2 Ease of preparation

Positive control materials can ideally accommodate different experimental setups (e.g., different solvents). For chemicals used in a liquid-based exposure system, the ease of preparation of a

stock suspension of a chemical in the solvent of interest can have a substantial impact on the stability and precision of the positive control material. For example, in the ongoing development of a skin sensitization assay at the National Institute of Standards and Technology, both 2,4-dinitrochlorobenzene (DNCB) and benzyl bromide were investigated as positive control materials. DNCB is challenging to weigh precisely because it is an oily solid, while the benzyl bromide powder is easier to weigh, which is a desirable trait for a positive control material. Ease of weighing can also be a challenge for engineered nanomaterials in powder form, where static forces can necessitate weighing the material first into an aluminum container prior to adding to a plastic or glass container and, even then, losses due to spilling from the static electricity need be carefully minimized. While these challenges do not preclude the use of a material as a positive control, the most straightforward of the relevant options should be selected to avoid potential complications.

Test method developers can consult guidance provided by other groups (e.g., US Pharmacopeia) related to the acceptable precision for weighing substances. To increase precision, a sufficient mass of the test substance should be weighed. For liquid substances, adding a precise volume can also be challenging for viscous substances or those that adhere to the pipette tip. Pipetting small volumes (e.g., < 10 µL) should be avoided due to reduced precision in transferring this volume range. In addition to the weighing process, challenges related to preparing the stock suspension include incomplete dissolution or the potential for reprecipitation. Dissolution is often evaluated by observing the solution for the presence of gradients or precipitates. A lack of complete dissolution may cause higher variability in the results because the substance in the delivery vehicle may not be homogeneous among different aliquots.

For some product types, such as medical devices, having a positive control material that can be dissolved in various solvents spanning the range of solvents that are used within the assay's applicability domain is advantageous. For cell-based assays in which the exposure occurs under submerged conditions, this will often include aqueous solvents, such as phosphate buffer, and potentially semi-polar solvents, such as dimethyl sulfoxide (DMSO). It may be necessary for some contexts of use, such as testing the extractions from medical devices (ISO, 2010) or consumer products (United States, 1973), to use polar, semi-polar, and non-polar solvents. It may be that a single positive control material is not dissolvable in this range of solvent types, in which case it may be necessary to have more than one positive control material for an assay. For testing medical devices, it may be possible to have a single positive control material, such as a polymer, that can undergo chemical extraction using different solvents (Nomura et al., 2018; Coleman et al., 2018).

4.3 Chemical purity

The performance of a positive control material can be impacted by its purity. Highly pure positive control materials produce less variability than less pure materials because impurities, which may vary between substance batches, may generate a biological response in the assay for the endpoint being evaluated or could cause interference (e.g., produce an absorbance or fluorescence signal similar to that being measured in the assay), impacting the assay result. Positive control materials may also be complex mixtures (e.g., biological extracts). As such, they may contain chemicals, some of which may be unknown, that elicit different mechanisms of action, which may problematic if assay developers aim to target a single mechanism of action. For example, uncertainty can arise from the use of positive control materials with a range of similar chemical structures (e.g., standard preparations of endotoxin, which are a mixture of lipopolysaccharides derived from *Escherichia coli*) or a product that is a complex mixture (e.g., fetal bovine serum). These issues can be alleviated by using chemically synthesized positive control materials that have a single chemical structure. If a positive control material with a chemically defined structure is not available, additional steps can ensure its quality. Endotoxin reference standards, for instance, are routinely prepared in international collaborative studies using the same bacterial strains, growth conditions, extraction methods, purification procedures, and chemical analyses (Rudbach et al., 1976). Considering the biological source and inherent heterogeneity of endotoxin batches produced in this manner, new batches are further defined by their biological activity relative to existing batches using compendial assays (Poole et al., 2012).

4.4 Verifiable physical properties

Physical properties of a positive control material, such as the identity, purity, concentration, and mass, should be easily measurable. These properties of reference materials should be evaluated using multiple independent analytical methods that are traceable to international standards (OECD, 2018). This reduces the risk of undetected impurities being present at sufficiently high concentrations to impact the positive control's dose-response function.

The selection of cadmium sulfate for an MTS (3-[4,5-dimethylthiazol-2-yl]-5-[3-carboxymethoxyphenyl]-2-[4-sulfophenyl]-2H-tetrazolium) nanocytotoxicity assay is a successful example of the use of a positive control material with known purity and concentration. The cadmium sulfate positive control material can be readily quantified using multiple methods, such as inductively coupled plasma-mass spectrometry, and the concentration is traceable to a reference material (e.g., NIST standard reference material (SRM) 3108 – Cadmium (Cd) Standard Solution) (Elliott et al., 2017; Rösslein et al., 2015). Results from an interlaboratory comparison of this assay revealed that the results for the positive control experiments of one laboratory were an outlier compared to the other four laboratories (Elliott et al., 2017). Because the cadmium concentration is readily quantifiable, it was straightforward to ensure that differences in the cadmium sulfate concentration in vials was not the cause of the discrepant results. Comparisons among the laboratories later indicated that unintentionally removing cells during the rinsing process by one laboratory caused the discrepant results.

4.5 Stability

Having a positive control material that is stable in storage and in a stock suspension facilitates performing the assay in a repeatable and reproducible manner. Good laboratory practice (GLP)

and good manufacturing practice (GMP) conditions require conducting stability studies. Stability studies should be performed for all compounds and solutions made, and should include information about freezing/thawing, storage (i.e., benchtop, refrigerator, or freezer), and any other conditions relevant to the study. These studies should take into account the amount of time the positive control material will be in any given storage condition. Separate freezing/thawing studies need to be conducted for any positive control material that will be frozen (OECD, 2018; US FDA, 2018). To evaluate the stability of a stock suspension, careful documentation is required regarding when stock suspensions are prepared, and the results for the stock suspension should be monitored over time to assess how long a particular positive control material yields a consistent result. Efficacy of the positive control material should remain within acceptability ranges given the storage and use conditions over a specified time frame. In general, positive control materials with longer stability are advantageous because fewer cross-validation studies will be needed to ensure a similar performance among batches of the material.

Instability of positive control materials during storage and handling may pose a fundamental challenge to the repeatability and reproducibility of the assay. Depending upon how unstable a positive control material is, it may be challenging to demonstrate a repeatable dose-response function. An example of positive control materials that degrade over time are those that generate reactive oxygen species (ROS) and reactive nitrogen species (RNS) (e.g., H_2O_2 or SIN-1) and need to be freshly prepared each time the assay is performed. ROS/RNS are often the cause of cellular damage, which may lead to downstream effects including acute toxicity, inflammation, or genotoxicity. Accurately correlating the concentration of ROS/RNS precursors with cell damage would be ideal, yet the inherent chemical instability of these precursors makes the accurate determination of their concentration very difficult (Roesslein et al., 2013). For example, the initial concentration of H_2O_2 in a stock solution is not precisely known. Operations, such as dilution steps or simple pipetting, can cause decomposition of the instable ROS/RNS precursors due to the increased exposure to O_2, severely limiting their capability as a positive control material (Roesslein et al., 2013). Even if the concentration of a ROS/RNS precursor can be analyzed through chemical analysis before performing the assay, it is challenging to ensure that this value remains accurate for the duration of the assay. Therefore, quantitative comparisons between ROS/RNS precursors are difficult (Roesslein et al., 2013), although qualitative evaluations using categories, such as mild, moderate, or severe effects, may be possible.

4.6 Generates responses spanning the assay dynamic range

Positive control experiments can be used to evaluate the repeatability and reproducibility of the assay dose-response function and to calibrate test samples. To evaluate repeatability and reproducibility, it is important for the positive control material to yield responses spanning the dose-response range. Measurements made at different concentrations ranging from no response to the maximum response capture the assay's full dynamic range and set the measurement bounds of the assay. Assay dose-response functions can often be fitted to linear or logistic models, and the fitting parameters can provide evidence that assays conducted on different days or in different laboratories are likely comparable; in the context of this paper, the assay dose-response function will refer to that of the positive control material. While the optimal number of concentrations to test for a dose-response curve for toxicological studies has been previously evaluated (Holland-Letz, 2017; Holland-Letz and Kopp-Schneider, 2015) and is beyond the scope of this paper, it is possible to establish the assay repeatability (e.g., that the instrumentation is performing consistently across the assay's dynamic range) and reproducibility by having a comprehensive understanding of the assay dose-response function.

A second function of a positive control experiment is to generate a calibration curve to enable comparisons among unknown experimental test samples. The implicit assumption is that the assay results for the test substance may be described in terms of the activity caused by a specific concentration of a positive control material. For example, this approach is used in *in vitro* endocrine disruption experiments, where 17β-estradiol or testosterone is the standard positive control material. Often, in assays designed to mimic a biological response, a positive control material "equivalent" score will be the assay result, such as "17β-estradiol equivalent," "methyl-testosterone equivalent", or "2,3,7,8-TCDD equivalent". This means that the compound at a given concentration has the same magnitude of a response in the assay as the positive control material at a given concentration (EPA, 2014). Other commonly used assay test material response readouts are expressed as a percentage of positive control response (at a designated active concentration). For this function of a positive control experiment, using complementary methods to measure the concentration of the positive control material will increase confidence in the generated calibration curve.

4.7 Technical and biological interference

Interference, in the context of this manuscript, describes the processes through which the positive control material interacts with other constituents in an assay or produces an aberrant signal (e.g., absorbance at a particular wavelength) similar to the intended measurand. Interferences can enhance or diminish a readout of an assay and bias the results. Potential interference for certain types of test substances has been observed in various *in vitro* assays. For example, the potential for engineered nanomaterials and other particulate substances to cause interferences or artifactual results is widely recognized (Hanna et al., 2018; Petersen et al., 2014, 2020; Keene et al., 2014; Guadagnini et al., 2015). It is also possible for other substances to cause technological interferences in an assay if they have an absorption or fluorescent signal similar to that of the probe molecule in an assay and are present at a sufficiently high concentration. Another possibility is that components of the cell culture medium, such as serum, could interact with the positive control material and change its biological capacity to generate the desired assay response. In addition, impurities (addressed above) may cause interferences in the assay response; these interferences are limited when highly pure positive control materials are used. Test-

ing for potential interference effects from a positive control material is assay-specific and should be evaluated as appropriate. Pharmacological counter-screens, such as ALAMR NMR, have been developed to identify promiscuous compounds, such as those causing nonspecific enzymatic inhibition and pan-assay interference compounds (PAINS) (Huth et al., 2005; Dahlin et al., 2017). Recently, structure-based computational tools[1] were developed to predict chemical-assay interference via luciferase inhibition and autofluorescence, using various cell types and culture conditions (Borrel et al., 2020a,b).

4.8 Commercial availability

Another key consideration for a positive control material is its long-term availability. The best scenario is for the material to be available through multiple commercial vendors, thus avoiding challenges if one vendor stops selling the material. Long-term availability of a positive control material is also supported when a robust synthesis procedure is publicly available, enabling any manufacturer to continue production.

To minimize variability, a positive control material would preferably be purchased from a single vendor for use within a laboratory or in an interlaboratory validation. An alternative focus is to evaluate the robustness of the positive control material by comparing results among vendors to evaluate what range of variability could be encountered by different groups performing an assay who may select different vendors. Any time a new lot or vendor is used by a laboratory, a cross validation study should be conducted to ensure the positive control materials are responding equivalently. In this instance, cross validation refers to ensuring that the positive control materials pass all quality control (QC) parameters in a given number of tests to establish confidence that an equivalent or known response is obtained in the new batch of positive control material. If the response of the new positive control material is significantly different or out of the QC range, a new set of QC parameters may need to be established for the new positive control material. Efforts should also be made to verify purity with each vendor. For example, a certificate of analysis from the vendor should document the specifications for a given batch or lot. In addition, expensive substances should be avoided for the positive control material, when possible, since this may limit the widespread adoption of an assay.

Anecdotally, one of the co-authors found a promising positive control material for an assay under development, but it was only available for purchase once a year at a mass of 10 mg per year per organization. Thus, if the bottle containing the positive control material was unintentionally damaged, that laboratory would have to wait up to a year before obtaining more of the substance. As another example, a lack of commercial availability of a reasonably priced proficiency chemical for Organisation for Economic Co-operation and Development (OECD) test guideline 456 (human chorionic gonadotropin) was recently noted and could hinder new laboratories from proving their proficiency with this assay (Kolle et al., 2019).

4.9 User toxicity

It is ideal for a positive control material to be of low toxicity to operators in case exposure occurs. It is also important to consider that many laboratories may not have high biosafety level certificates, and thus these types of positive control materials should be avoided if possible. It is also worth considering the types of laboratories that may be conducting a certain test, as academic settings with graduate students may be less willing to use more toxic substances than a contract research organization. For example, one of the co-authors was involved in a study where a decision was made not to use asbestos as a positive control material for a respiratory toxicity assay to avoid potential exposure to graduate students at a university conducting the research.

It is important to note that the toxicity of the positive control material also depends upon the concentration that is needed to cause the desired response in the assay; using a high concentration of a chemical may raise the risk to workers for a chemical that may otherwise be perceived as being of relatively low toxicity. For example, a higher concentration may be needed when exposing more robust cell types or if the deposited dose is only a small fraction of the generated dose (e.g., for exposure of cells grown at the air-liquid interface to an aerosolized substance); therefore, certain assays may require a sufficiently high concentration that could pose worker health risks.

4.10 Disposability

Ideally, waste disposal should be straightforward. The positive control material should not be likely to react with other compounds in the waste disposal collection unit or require expensive disposal. For assays evaluating toxicity endpoints, the positive control materials will also likely be toxic to some degree, and thus there may be specific disposal requirements based on the type of compound (e.g., pesticide or heavy metal). On the other hand, if an antibiotic is used as a positive control material, it can be degraded at high temperatures prior to disposal to avoid disposal as a toxic material.

5 Case study evaluating the use of lipoteichoic acid as a positive control material for the monocyte activation test

The safe use of drugs, medical devices, and other medical products relies on accurate detection of contaminating agents. Pyrogens are contaminants that include constituents from gram-positive or gram-negative bacteria, viruses, fungi, or other substances, which cause an inflammatory response leading to a fever response in humans (Netea et al., 2000). Most regulated medical products must not contain pyrogen levels above specified limits, and several test methods can be used to quantify pyrogen contamination to different levels of accuracy and precision. Currently, pyrogen testing is often conducted using the rabbit pyrogen test, which requires at least three rabbits per test. Each year, an estimated 130,000 rabbit pyrogen tests are conducted worldwide,

[1] https://sandbox.ntp.niehs.nih.gov/interferences/

requiring approximately 400,000 rabbits (Hartung, 2015). These numbers are loose extrapolations, since in most regions there are no requirements for companies or regulators to track or disclose the number of animals used in specific procedures. This is not the case in Europe, though, where these figures are routinely collected. In Europe, roughly 35,000 rabbits were used for pyrogen tests in 2017 (EC, 2020).

Several *in vitro* methods have been developed to assess pyrogen contamination. The monocyte activation test (MAT) detects the innate immune response in human monocytes by quantifying the degree of pro-inflammatory cytokine release following exposure to substances that induce fever (Hartung, 2015; Molenaar-de Backer et al., 2021). The MAT can assess the presence of diverse types of pyrogenic substances, and there are several positive control materials that can be used to show monocyte activation.

The MAT uses human whole blood, peripheral blood mononuclear cells (PBMCs), or the Mono-Mac-6 (MM6) cell line (Hartung, 2015, 2021). Regardless of cell source, monocytes recognize pathogens through toll-like receptors (TLRs), whose signaling cascade ultimately leads to the pro-inflammatory cytokine release that is the hallmark of the human fever response (Hasiwa et al., 2013; Netea et al., 2000). Various factors can trigger the TLR response and be detected by the MAT, including general bacterial constituents, flagellin, double-stranded RNA, and fungal constituents, but bacteria are the most common contaminant of greatest concern (Hayashi et al., 2001; Alexopoulou et al., 2001; Figueiredo et al., 2011; Borton and Coleman, 2018). One poorly defined group of materials, referred to as "material mediated pyrogens" (MMPs), are thought to cause fever, but there is no literature available to support this claim (Borton and Coleman, 2018). Furthermore, while MMPs have been suggested to have pyrogenic effects *in vivo*, there is no consensus that these materials mechanistically are pyrogens or that they are likely to be involved in the medical device manufacturing process (e.g., cocaine and LSD are two examples given in the ISO 10993 document, both of which are described as disruptors of neurological thermoregulatory centers rather than activators of the inflammatory signaling cascade) (Borton and Coleman, 2018). Substances that are cytotoxic to blood cells interfere with the test system and may not be compatible with this method (ICCVAM, 2008).

Heat-killed gram-negative or gram-positive bacteria contain complex bacterial constituents that can trigger diverse TLRs, and the MAT is broadly compatible with a wide range of substances and materials used in medical products. For gram-negative bacteria, lipopolysaccharide (LPS, endotoxin) is a ligand for TLR4 and represents the most significant class of pyrogen for most biomedical products (Park and Lee, 2013). For gram-positive bacteria, major cell wall components, such as peptidoglycan and lipoteichoic acid (LTA), trigger TLR2 receptors (Schwandner et al., 1999). Overall, possible positive control materials for the MAT will test for monocyte activation through TLR-mediated signal transduction.

The selection of a positive control material can depend on the regulation and the purpose of testing. Because the vast majority of contaminating pyrogens are endotoxins derived from gram-negative bacteria, the international reference standard endotoxin (RSE) is the reference material commonly accepted by the World Health Organization, US Pharmacopeia, European Pharmacopoeia, and Japanese Pharmacopoeia (US Pharmacopeia, 2012; Poole et al., 2012). However, there are no internationally harmonized reference materials for gram-positive bacteria, and there are differences in guidance for the use of non-endotoxin positive control materials among governing bodies. Guidance from the US FDA (US FDA, 2012, 2016) refers to the International Organization for Standardization (ISO, 2012). ISO 10993-12 specifies that reference materials should be validated by individual laboratories, but provides no guidance on the selection of reference materials (ISO, 2012, section 5.1). Furthermore, positive control materials are recommended to display grades of responses, such as minimum, intermediate, and severe (ISO, 2012, section 6), a recommendation in agreement with characteristic number 6 ("Generates responses spanning the assay dynamic range") described above. The European Pharmacopoeia specifies that two non-endotoxin positive control materials should be used once they are qualified, including peptidoglycans, lipoteichoic acid (LTA), synthetic bacterial lipoproteins, flagellin, or crude bacterial whole cell extract (EDQM, 2020). The recommendation of multiple positive control materials with different mechanisms of action, all of which are relevant to the outcome being measured in the assay, is reflective of characteristic number 1 ("Biological mechanism of action") described above. To be qualified, the positive control material must follow a significant log-linear response (characteristic number 6 described above) and have the ability to detect previous contaminants or match previous methodologies (EDQM, 2020). For a quantitative approach, a standard curve must be performed at four doses to estimate potential pyrogen levels, which falls under characteristic number 6 described above. Pass/fail testing of pyrogenicity can be determined by comparing the test article to a threshold (contaminant limit concentration).

One of the key advantages of the MAT is its ability to detect gram-positive bacteria (Hasiwa et al., 2013). It is important to have a gram-positive bacterial control so that the assay is confirmed to function as expected with this mechanism. Below, we present a case study evaluating the characteristics and appropriateness of using LTA as a positive control material in the MAT for pyrogen testing of medical devices. Medical devices represent a diverse set of products that come into contact with bodily fluids through implantation or injection; therefore, MAT assay conditions can vary depending on the device shape, size, or material make-up (Brown, 2021). Here, we consider use of LTA as a positive control material for different types of medical devices and assay formats and evaluate it based on the ten characteristics discussed above.

1. Biological mechanism of action

LTA is a cell wall polymer from gram-positive bacteria that can be naturally- or synthetically-derived. A comparison of LTA from these two sources can be found in Table 1; LTA from both sources has been mechanistically well characterized. Human monocytes sense LTA through TLR2 receptors and trigger a signal cascade to release pro-inflammatory cytokines, such as IL-6, IL-1β, and TNF-α (Schwandner et al., 1999; Netea et al., 2000). An enzyme-

Tab. 1: Outline of the characteristics of two sources of lipoteichoic acid (LTA): 1) naturally-derived and 2) chemically-defined
Each source has advantages and disadvantages to its use; however, both are suitable positive control materials for the monocyte activation test (MAT).

Characteristic	Naturally-derived LTA	Chemically-defined LTA
1. Biological mechanism of action	The mechanism of action of naturally-derived LTA is well known and it reflects the relevant mechanism of action being studied in the MAT (TLR2 activation resulting in cytokine release from monocytes; Stadelmaier et al., 2003).	Studies have elucidated the minimal structure and chain length needed for the successful use of chemically-defined LTA as a positive control material in the MAT (through TLR2 activation; Deininger et al., 2007, 2003).
2. Ease of preparation	LTA is a component of the cell wall of gram-positive bacteria. It is easy to use LTA in the form of heat-killed bacteria or to purify it from bacteria (de Oliviera Nascimento et al., 2012).	LTA is synthesized in a chemical process. The chemical structures and mixture ratios are quantified (Morath et al., 2002; Stadelmaier et al., 2003).
3. Chemical purity	LTA is amphipathic and can be dissolved in many polar (e.g., aqueous) and non-polar solvents (Morath et al., 2005). As a complex mixture and natural product, the constituents may vary between strains and vendors.	Chemically-defined LTA possesses the same amphipathic properties as naturally-derived LTA and can dissolve in polar (e.g., aqueous) and non-polar solvents (Morath et al., 2005). A highly purified LTA can be produced.
4. Verifiable physical properties	NMR and LC-MS can be used to quantify the identity, purity, and concentration of the material (Morath et al., 2005). Other methods can be used to ensure preparations are endotoxin-free.	
5. Stability	Naturally-derived LTA is stable for 1-6 months, depending on storage temperature and handling. It contains no chemical groups that can undergo redox or other modifications.	Chemically-defined LTA is a smaller polymer of the naturally-derived LTA, and its stability is expected to be similar (Morath et al., 2005).
6. Generates responses spanning the assay dynamic range	Both can be used at concentrations that do not initiate a response (minimum readout value) and concentrations that demonstrate a saturated response (maximum assay readout value; Deininger et al., 2003).	
7. Technical and biological interference	Neither shows evidence of enhancing or diminishing the assay read-out (i.e., cytokine release) through either technological interference or indirect biological activity.	
8. Commercial availability	Naturally-derived, endotoxin-free LTA is commercially available at reasonable cost. Vendors have certificates of analysis.	Chemically-defined LTA is not commercially available, but the synthesis process is published (Stadelmaier et al., 2003).
9. User toxicity	There is low expected toxicity for operators given that the correct personal protective equipment is used.	
10. Disposability	Both can be disposed through a professional waste disposal service without additional complications.	

linked immunosorbent assay (ELISA) quantifies cytokine release, preferably using sequence-defined recombinant antibodies (Groff et al., 2020). LTA can be used to evaluate monocyte activation in response to gram-positive bacteria.

2. Ease of preparation
LTA can be prepared in the desired matrix for pyrogen testing of medical devices. In order to test a device eluate, medical devices can be spiked with the positive control material followed by extraction in polar (e.g., PBS) or non-polar (e.g., sesame oil) vehicles (ISO, 2012, section 10.3.5). LTA is an amphipathic substance that can be extracted using polar or non-polar solvents (Morath et al., 2005). In addition, the MAT allows for smaller medical devices to be tested by complete immersion in human whole blood, PBMCs, or MM6 cells (ISO, 2012, section 10.3.3), avoiding the additional complexity of an extraction step, which itself needs a positive control material to ensure it was performed successfully. Thus, LTA is amenable to diverse assay formats and a range of solvents.

3. Chemical purity
LTA of varying purity may be used, which will affect the performance of the MAT assay. For example, naturally-derived LTA may be used as heat-killed bacteria or may be extracted from bacteria. Heat-killed gram-positive bacteria possess LTA and other pyrogenic constituents recognized by monocytes (de Oliviera Nascimento et al., 2012). Alternately, LTA can be extracted and purified from bacteria, and its immunostimulatory ability will depend on the extraction method. There is greater monocyte activation

using LTA extracted with butanol rather than phenol (Morath et al., 2005). Contamination with endotoxin (which is of particular concern when using higher LTA concentrations) should be avoided in LTA preparations. The absence of endotoxin can be confirmed using commercially available kits (e.g., HEK-Blue TLR4 cells (Invivogen, USA) or recombinant factor C assay (Lonza, Switzerland)). Whether heat-killed or extracted, there may be differences depending on the strain of bacteria used, which may influence interlaboratory comparisons.

In contrast to naturally-derived sources, LTA can be synthesized as a shorter polymer with immunostimulatory ability similar to the natural form (Deininger et al., 2003, 2007; Stadelmaier et al., 2003). While natural LTA allows for broad monocyte recognition, the synthetic version allows more control over the production of the reference material.

4. Verifiable physical properties

Physical properties, such as the identity, purity, concentration, and mass, should be assessed before using LTA. There are multiple methods for assessing the purity of LTA. LTA extractions have been evaluated with hydrophobic interaction chromatography, followed by mass spectroscopy (Morath et al., 2005). However, chromatography measurements are destructive. NMR offers a non-destructive LTA characterization method (Morath et al., 2002). While not a physical property of LTA, these assessments could also uncover contaminants, such as endotoxin (LPS) contamination.

5. Stability

The stability of LTA is critical for determining its usefulness as a positive control material. Information is available for the natural form of purified LTA. Upon suspension, purified LTA from *S. aureus* can be stored at 4°C for 1 month and -20°C for 6 months (Invivogen, USA). It is recommended to avoid freeze-thaw cycles (Invivogen, USA). Stability information is lacking for synthetic LTA; however, structural quality, such as polymer length, can be monitored over time with NMR and mass spectroscopy.

6. Generates responses spanning the assay dynamic range

Acceptance criteria for a valid positive control MAT experiment are based on the properties of the dose-response function. The European Pharmacopoeia outlines two critical criteria: 1) a log-linear relationship, and 2) a statistically significant slope in the regression (EDQM, 2020). These two criteria can be historically tracked and can form the basis of comparison for monitoring consistent among-assay performance.

7. Technical and biological interference

There are no known reports of LTA enhancing or diminishing the read-out of cytokine concentration in the ELISA assay. Interference can be assessed by measuring cytokine release in the presence of LTA in the absence of cells. Separately, it is possible for medical devices themselves to cause interference with the ELISA cytokine read out. As per the European Pharmacopoeia, medical device extracts are tested within the ELISA detection system to account for the effect of interference (EDQM, 2020).

8. Commercial availability

The availability of highly pure LTA is critical for widespread use in the MAT. Natural LTA is accessible at multiple commercial vendors (e.g., Sigma or Invivogen) at high purity, at a reasonable cost, and without detectable endotoxin contamination. Highly pure synthetic LTA is not currently commercially available, but the synthesis process is published (Morath et al., 2002), which will facilitate commercialization.

9. User toxicity and 10. disposability

LTA is relatively safe to handle and dispose. *In silico* Collaborative Acute Toxicity Modeling Suite (CATMoS) models (Kleinstreuer et al., 2018) implemented in OPERA (version 2.5[2]) predicts LTA (CASRN: 56411-57-5; PubChem) to have low acute oral toxicity (GHS category 5). The Safety Data Sheet (Invivogen, USA) of a natural LTA directs users to avoid contact with eyes, skin, and repeated exposure. While respiratory protection is not required, inhalation should be avoided and personal protective equipment, such as impermeable gloves and lab coat, should be worn. According to the Safety Data Sheet (Invivogen, USA), a licensed professional waste disposal service should be in charge of disposing LTA.

Summary on the use of LTA as a positive control material in the MAT

Natural or synthetic LTA shows promise as a suitable positive control material based on the aforementioned set of characteristics. Its mechanism of action is well known, and it is easy to prepare, use, and dispose of. Multiple, independent methods are able to assess the purity of LTA. Naturally-derived and synthetic LTA can elicit similar magnitudes of effect in the MAT. Naturally-derived LTA isolated from gram-positive bacteria is commercially available at a reasonable cost and can be used in currently available MAT kits from numerous vendors. Synthetic LTA is not currently commercially available, but the process to synthesize LTA is published and is publicly available to use for commercialization. Future efforts should focus on developing a commercially available synthetic LTA and showing its reproducibility in the MAT. In addition to LTA, other positive control materials should be considered, especially considering that the European Pharmacopoeia recommends two non-endotoxin pyrogen positive control materials. For example, general bacterial components, such as triacylated lipoproteins (Aliprantis et al., 1999) and CpG oligonucleotides (Hartmann and Krieg, 1999), capture both gram-positive and gram-negative pyrogen sources, and are therefore good candidates for consideration under the ten characteristics outlined here. Establishing a harmonized positive control material for non-endotoxin pyrogens will build greater confidence in the MAT and lead to more widespread adoption.

[2] https://github.com/kmansouri/OPERA

6 Conclusion

The characteristics outlined in this paper will help in selecting potential positive control materials. The selection of a positive control material involves consideration of the following: 1) the biological mechanism of action, 2) ease of preparation, 3) chemical purity, 4) verifiable physical properties, 5) stability, 6) ability to generate responses spanning the dynamic range of the assay, 7) technical and biological interference, 8) commercial availability, 9) user toxicity, and 10) disposability. *In vitro* assay developers should recognize that these characteristics are goals to consider rather than absolute requirements. Fulfilling these characteristics is ideal for the longevity and robustness of an assay, but selection of a positive control material may require flexibility in how stringently they are applied.

References

Alexopoulou, L., Holt, A. C., Medzhitov, R. et al. (2001). Recognition of double-stranded RNA and activation of NF-κB by toll-like receptor 3. *Nature 413*, 732-738. doi:10.1038/35099560

Aliprantis, A. O., Yang, R. B., Mark, M. R. et al. (1999). Cell activation and apoptosis by bacterial lipoproteins through toll-like receptor-2. *Science 285*, 736-739. doi:10.1126/science.285.5428.736

Avila, A. M., Bebenek, I., Bonzo, J. A. et al. (2020). An FDA/CDER perspective on nonclinical testing strategies: Classical toxicology approaches and new approach methodologies (NAMs). *Regul Toxicol Pharmacol 114*, 104662. doi:10.1016/j.yrtph.2020.104662

Barosova, H., Maione, A. G., Septiadi, D. et al. (2020). Use of Epi-Alveolar lung model to predict fibrotic potential of multiwalled carbon nanotubes. *ACS Nano 14*, 3941-3956. doi:10.1021/acsnano.9b06860

Basketter, D., Azam, P., Casati, S. et al. (2019). Applying non-animal strategies for assessing skin sensitisation report from an EPAA/cefic-LRI/IFRA Europe cross sector workshop, ECHA, February 7th and 8th 2019. *Regul Toxicol Pharmacol 109*, 104477. doi:10.1016/j.yrtph.2019.104477

Borrel, A., Huang, R., Sakamuru, S. et al. (2020a). High-throughput screening to predict chemical-assay interference. *Sci Rep 10*, 3986. doi:10.1038/s41598-020-60747-3

Borrel, A., Mansouri, K., Nolte, S. et al. (2020b). InterPred: A webtool to predict chemical autofluorescence and luminescence interference. *Nucleic Acids Res 48*, W586-W590. doi:10.1093/nar/gkaa378

Borton, L. K. and Coleman, K. P. (2018). Material-mediated pyrogens in medical devices: Applicability of the in vitro monocyte activation test. *ALTEX 35*, 453-463. doi:10.14573/altex.1709221

Brown, J., Clippinger, A. J., Fritz Briglia, C. et al. (2021). Using the monocyte activation test as a stand-alone release test for medical devices. *ALTEX 38*, 151-156. doi:10.14573/altex.2012021

Browne, P., Judson, R. S., Casey, W. M. et al. (2015). Screening chemicals for estrogen receptor bioactivity using a computational model. *Environ Sci Technol 49*, 8804-8814. doi:10.1021/acs.est.5b02641

Chipinda, I., Ajibola, R. O., Morakinyo, M. K. et al. (2010). Rapid and simple kinetics screening assay for electrophilic dermal sensitizers using nitrobenzenethiol. *Chem Res Toxicol 23*, 918-925. doi:10.1021/tx100003w

Coleman, K. P., Grailer, T. P., McNamara, L. R. et al. (2018). Preparation of irritant polymer samples for an in vitro round robin study. *Toxicol In Vitro 50*, 401-406. doi:10.1016/j.tiv.2018.01.018

Dahlin, J. L., Nelson, K. M., Strasser, J. M. et al. (2017). Assay interference and off-target liabilities of reported histone acetyltransferase inhibitors. *Nat Commun 8*, 1527. doi:10.1038/s41467-017-01657-3

de Oliviera Nascimento, L., Massari, P. and Wetzler, L. (2012). The role of TLR2 in infection and immunity. *Front Immunol 3*, 79. doi:10.3389/fimmu.2012.00079

Deininger, S., Stadelmaier, A., von Aulock, S. et al. (2003). Definition of structural prerequisites for lipoteichoic acid-inducible cytokine induction by synthetic derivatives. *J Immunol 170*, 4134-4138. doi:10.4049/jimmunol.170.8.4134

Deininger, S., Figueroa-Perez, I., Sigel, S. et al. (2007). Use of synthetic derivatives to determine the minimal active structure of cytokine-inducing lipoteichoic acid. *Clin Vaccine Immunol 14*, 1629-1633. doi:10.1128/CVI.00007-07

EC – European Commission (2020). 2019 Report on the Statistics on the Use of Animals for Scientific Purposes in the Member States of the European Union in 2015-2017. https://ec.europa.eu/environment/chemicals/lab_animals/pdf/Statistical%20report_SWD_p1_EN.pdf

EDQM – European Directorate for the Quality of Medicines and HealthCare (2020). Monocyte Activation Test Chapter 2.6.30 European Pharmacopoeia (Version 07/2017), 10th edition.

Elliott, J. T., Rosslein, M., Song, N. W. et al. (2017). Toward achieving harmonization in a nanocytotoxicity assay measurement through an interlaboratory comparison study. *ALTEX 34*, 201-218. doi:10.14573/altex.1605021

EPA – Environmental Protection Agency (2014). SW-846 Test Method 4435: Screening for Dioxin-Like Chemical Activity in Soils and Sediments Using the Calux Bioassay and Toxic Equivalents (TEQs) Determinations. https://bit.ly/3e83cr1

Figueiredo, R. T., Carneiro, L. A. M. and Bozza, M. T. (2011). Fungal surface and innate immune recognition of filamentous fungi. *Front Microbiol 2*, 248. doi:10.3389/fmicb.2011.00248

Groff, K., Allen, D., Casey, W. et al. (2020). Increasing the use of animal-free recombinant antibodies. *ALTEX 37*, 309-311. doi:10.14573/altex.2001071

Guadagnini, R., Kenzaoui, B. H., Walker, L. et al. (2015). Toxicity screenings of nanomaterials: Challenges due to interference with assay processes and components of classic in vitro tests. *Nanotoxicology 9*, 13-24. doi:10.3109/17435390.2013.829590

Hanna, S. K., Montoro Bustos, A. R., Peterson, A. W. et al. (2018). Agglomeration of Escherichia coli with positively charged nanoparticles can lead to artifacts in a standard Caenorhabditis elegans toxicity assay. *Environ Sci Technol 52*, 5968-5978. doi:10.1021/acs.est.7b06099

Hartmann, G. and Krieg, A. M. (1999). CpG DNA and LPS induce distinct patterns of activation in human monocytes. *Gene Ther 6*, 893-903. doi:10.1038/sj.gt.3300880

Hartung, T., Balls, M., Bardouille, C. et al. (2002). Good cell culture practice. ECVAM good cell culture practice task force report 1. *Altern Lab Anim 30*, 407-414. doi:10.1177/026119290203000404

Hartung, T. (2015). The human whole blood pyrogen test – Lessons learned in twenty years. *ALTEX 32*, 79-100. doi:10.14573/altex.1503241

Hartung, T. (2021). Pyrogen testing revisited on occasion of the 25th anniversary of the whole blood monocyte activation test. *ALTEX 38*, 3-19. doi:10.14573/altex.2101051

Hasiwa, N., Daneshian, M., Bruegger, P. et al. (2013). Evidence for the detection of non-endotoxin pyrogens by the whole blood monocyte activation test. *ALTEX 30*, 169-208. doi:10.14573/altex.2013.2.169

Hayashi, F., Smith, K. D., Ozinsky, A. et al. (2001). The innate immune response to bacterial flagellin is mediated by toll-like receptor 5. *Nature 410*, 1099-1103. doi:10.1038/35074106

Holland-Letz, T. and Kopp-Schneider, A. (2015). Optimal experimental designs for dose-response studies with continuous endpoints. *Arch Toxicol 89*, 2059-2068. doi:10.1007/s00204-014-1335-2

Holland-Letz, T. (2017). On the combination of c- and D-optimal designs: General approaches and applications in dose-response studies. *Biometrics 73*, 206-213. doi:10.1111/biom.12545

Huth, J. R., Mendoza, R., Olejniczak, E. T. et al. (2005). ALARM NMR: A rapid and robust experimental method to detect reactive false positives in biochemical screens. *J Am Chem Soc 127*, 217-224. doi:10.1021/ja0455547

ICCVAM – Interagency Coordinating Committee on the Validation of Alternative Methods (2008). ICCVAM Test Method Evaluation Report: Validation Status of Five In Vitro Test Methods Proposed for Assessing Potential Pyrogenicity of Pharmaceuticals and Other Products.

ISO – International Organization for Standardization (2010). 10993-10:2010: Biological Evaluation of Medical Devices – Part 10: Tests for Irritation and Skin Sensitization. Geneva, Switzerland: International Organization for Standardization.

ISO (2012). 10993-12:2012: Biological Evaluation of Medical Devices – Part 12: Sample Preparation and Reference Materials. Geneva, Switzerland: International Organization for Standardization.

ISO (2015). Guide 30: 2015: Reference Materials – Selected Terms and Definitions. Geneva, Switzerland: International Organization for Standardization.

Keene, A. M., Bancos, S. and Tyner, K. M. (2014). Considerations for in vitro nanotoxicity testing. In S. C. Sahu and D. A. Casciano (eds.), *Handbook of Nanotoxicology, Nanomedicine and Stem Cell Use in Toxicology*. Chapter 2. John Wiley and Sons. doi:10.1002/9781118856017.ch2

Kleinstreuer, N. C., Ceger, P., Watt, E. D. et al. (2017). Development and validation of a computational model for androgen receptor activity. *Chem Res Toxicol 30*, 946-964. doi:10.1021/acs.chemrestox.6b00347

Kleinstreuer, N. C., Karmaus, A., Mansouri, K. et al. (2018). Predictive models for acute oral systemic toxicity: A workshop to bridge the gap from research to regulation. *Comput Toxicol 8*, 21-24. doi:10.1016/j.comtox.2018.08.002

Kolle, S. N., Hil, E., Raabe, H. et al. (2019). Regarding the references for reference chemicals of alternative methods. *Toxicol In Vitro 57*, 48-53. doi:10.1016/j.tiv.2019.02.007

Leibrock, L., Jungnickel, H., Tentschert, J. et al. (2020). Parametric optimization of an air-liquid interface system for flow-through inhalation exposure to nanoparticles: Assessing dosimetry and intracellular uptake of CeO_2 nanoparticles. *Nanomaterials 10*, 2369. doi:10.3390/nano10122369

Molenaar-de Backer, M. W. A., Gitz, E., Dieker, M. et al. (2021). Performance of monocyte activation test supplemented with human serum compared to fetal bovine serum. *ALTEX 38*, 307-315. doi:10.14573/altex.2008261

Morath, S., Stadelmaier, A., Geyer, A. et al. (2002). Synthetic lipoteichoic acid from Staphylococcus aureus is a potent stimulus of cytokine release. *J Exp Med 195*, 1635-1640. doi:10.1084/jem.20020322

Morath, S., von Aulock, S. and Hartung, T. (2005). Structure/function relationships of lipoteichoic acids. *J Endotoxin Res 11*, 348-356. doi:10.1179/096805105X67328

Nelson, B. C., Petersen, E. J., Marquis, B. J. et al. (2013). NIST gold nanoparticle reference materials do not induce oxidative DNA damage. *Nanotoxicology 7*, 21-29. doi:10.3109/17435390.2011.626537

Netea, M. G., Kullberg, B. J. and Van der Meer, J. W. M. (2000). Circulating cytokines as mediators of fever. *Clin Infect Dis 31*, *Suppl 5*, S178-S184. doi:10.1086/317513

Nomura, Y., Lee, M., Fukui, C. et al. (2018). Proof of concept testing of a positive reference material for in vivo and in vitro skin irritation testing. *J Biomed Mater Res B Appl Biomater 106*, 2807-2814. doi:10.1002/jbm.b.34061

OECD – Organization for Economic Cooperation and Development (2005). Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. *Series on Testing and Assessment, No. 34*. OECD Publishing, Paris. http://www.oecd.org/official documents/displaydocument/?cote=env/jm/mono(2005)14&doclanguage=en

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

OECD (2019). Test No. 442C: In Chemico Skin Sensitisation: Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins. *OECD Guidelines for the Testing of Chemicals, Section 4*. OECD Publishing, Paris. doi:10.1787/9789264229709-en

Park, B. S. and Lee, J. O. (2013). Recognition of lipopolysaccharide pattern by TLR4 complexes. *Exp Mol Med 45*, e66. doi:10.1038/emm.2013.97

Petersen, E. J., Henry, T. B., Zhao, J. et al. (2014). Identification and avoidance of potential artifacts and misinterpretations in nanomaterial ecotoxicity measurements. *Environ Sci Technol 48*, 4226-4246. doi:10.1021/es4052999

Petersen, E. J., Hirsch, C., Elliott, J. T. et al. (2020). Cause-and-effect analysis as a tool to improve the reproducibility of nanobioassays: Four case studies. *Chem Res Toxicol 33*, 1039-1054. doi:10.1021/acs.chemrestox.9b00165

Plant, A. L., Locascio, L. E., May, W. E. et al. (2014). Improved reproducibility by assuring confidence in measurements in biomedical research. *Nat Methods 11*, 895-898. doi:10.1038/nmeth.3076

Poole, S., Desai, T., Findlay, L. et al. (2012). WHO international standard for endotoxin: Report of an international collaborative study to evaluate three preparations of endotoxin for their suitability to serve as the third international standard for bacterial endotoxin. *World Health Organization*, 1-40. https://apps.who.int/iris/handle/10665/96606

Roesslein, M., Hirsch, C., Kaiser, J. P. et al. (2013). Comparability of in vitro tests for bioactive nanoparticles: A common assay to detect reactive oxygen species as an example. *Int J Mol Sci 14*, 24320-24337. doi:10.3390/ijms141224320

Rösslein, M., Elliott, J. T., Salit, M. et al. (2015). Use of cause-and-effect analysis to design a high-quality nanocytotoxicology assay. *Chem Res Toxicol 28*, 21-30. doi:10.1021/tx500327y

Rudbach, J. A., Akiya, F. I., Elin, R. J. et al. (1976). Preparation and properties of a national reference endotoxin. *J Clin Microbiol 3*, 21-25.

Schwandner, R., Dziarski, R., Wesche, H. et al. (1999). Peptidoglycan- and lipoteichoic acid-induced cell activation is mediated by toll-like receptor 2. *J Biol Chem 274*, 17406-17409. doi:10.1074/jbc.274.25.17406

Stadelmaier, A., Morath, S., Hartung, T. et al. (2003). Synthesis of the first fully active lipoteichoic acid. *Angew Chem Int Edit 42*, 916-920. doi:10.1002/anie.200390243

Urbisch, D., Mehling, A., Guth, K. et al. (2015). Assessing skin sensitization hazard in mice and men using non-animal test methods. *Regul Toxicol Pharmacol 71*, 337-351. doi:10.1016/j.yrtph.2014.12.008

United States (1973). Federal Hazardous Substances Act. Title 16. Code of Federal Regulations. Part 1500: United States Code.

US FDA – United States Food and Drug Administration (2012). Guidance for Industry. Pyrogen and Endotoxins Testing: Questions and Answers. https://www.fda.gov/media/83477/download

US FDA (2016). Use of International Standard ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process". Department of Health and Human Services Food and Drug Administration. https://www.fda.gov/media/85865/download

US FDA (2018). Bioanalytical Method Validation Guidance for Industry. https://bit.ly/386YXrZ

US Pharmacopeia (2012). Bacterial Endotoxins Test – Chapter 85. United States Pharmacopeia 37, 1-5. https://bit.ly/306sRIB

Conflict of interest

The authors declare that they have no conflicts of interest.