Cause-and-effect analysis as a tool to improve the reproducibility of nanobioassays: four case studies

Elijah J. Petersen*[1], Cordula Hirsch[2], John T. Elliott[1], Harald F. Krug[3], Leonie Aengenheister[2], Ali Talib Arif[4, 5], Alessia Bogni[6], Agnieszka Kinsner-Ovaskainen[6], Sarah May[2], Tobias Walser[7], Peter Wick[2], Matthias Roesslein[2]

1) Biosystems and Biomaterials Division, Material Measurement Laboratory, National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899

2) Empa, Swiss Federal Laboratories for Material Testing and Research, Particles-Biology Interactions Laboratory, CH-9014 St. Gallen, Switzerland

3) NanoCASE GmbH, St. Gallerstr. 58, 9032 Engelburg, Switzerland

4) Institute for Infection Prevention and Hospital Epidemiology, University Medical Center Freiburg, University of Freiburg, Faculty of Medicine, University of Freiburg, D- 79106 Freiburg, Germany

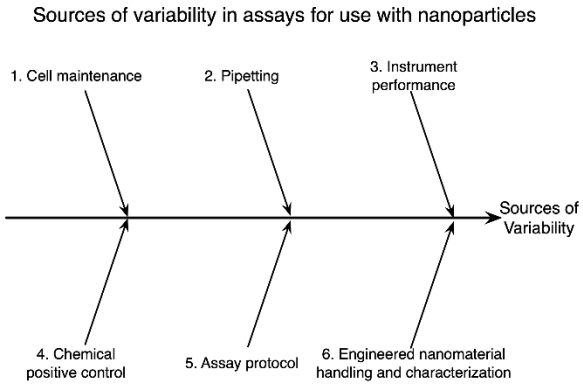
5) Kurdistan Institution for Strategic Studies and Scientific Research (KISSR), Qirga - Sulaimani, Iraq

6) European Commission, DG Joint Research Centre, 21027 Ispra, VA, Italy

7) Vereala GmbH, Eyhof 34, 8047 Zürich, Switzerland

***Corresponding author: E.J.P., elijah.petersen@nist.gov, Tel: (301)-975-8142**

TOC graphic



Abstract

One of the challenges in using *in vitro* data to understand the potential risks of engineered nanomaterials (ENMs) is that results often differ or are even contradictory among studies. While it is recognized that numerous factors can influence results produced by nanobioassays, there has not yet been a consistently used conceptual framework to identify key sources of variability in these assays. In this paper, we use cause-and-effect analysis to systematically describe sources of variability in four key *in vitro* nanobioassays: the DCF (2',7'-dichlorofluorescein) assay, an enzyme-linked immunosorbent assay (ELISA) for measuring interleukin-8, a flow cytometry assay (Annexin V/Propidium Iodide), and the Comet assay. These assays measure endpoints that can occur in cells impacted by ENMs through oxidative stress, a principle mechanism for ENM toxicity. The results from this analysis identify control measurements to test for potential artifacts or biases that could occur during conduct of these assays with ENMs. Cause-and-effect analysis also reveal additional measurements that could be performed either in preliminary experiments or each time the assay is run to increase confidence in the assay results and their reproducibility within and among laboratories. The approach applied here with these four assays can be used the support the development of a broad range of nanobioassays.

Introduction

Engineered nanomaterials (ENMs) often have unique or enhanced properties such as high surface reactivity and quantum confinement compared to bulk materials of the same elemental composition. These properties can be utilized in a broad range of commercial applications such as environmental remediation, biomedical application, textiles, and renewable energy.[1-5] However, ENMs may also pose potential risks to human health or the environment as a result of these same properties since a fraction of the ENMs will be intentionally or unintentionally released during the production and life cycle of these products.[6-15] Many assays used to assess biological effects of dissolved chemicals may require modifications prior to use with ENMs because these materials often behave differently in the test media (e.g., agglomeration or dissolution) and may cause artifacts in the test results.[16-22]

The most frequently used tests to assess the potential hazards of ENMs are *in vitro* toxicity assays.[17] They have the advantage of enabling higher throughput testing which is not possible when conducting tests with larger and more complex organisms (e.g., rats or fish). High throughput testing can be used to develop categories for ENMs based on results from these assays, the ENM composition, and physicochemical properties.[23-26] In some cases, there is evidence that *in vitro* assay results might also be used to estimate the toxicity observed with more complex organisms such as pulmonary toxicity with rats.[27,28] In addition, *in vitro* tests such as cytotoxicity assays can be used in a tiered testing system to assess potential hazards caused by ENMs. This can support the prioritization of additional testing using larger organisms thus reducing the number of animals used.[27-29]

One of the most substantial limitations for *in vitro* testing with ENMs is a lack of robust, "validated" methods for these materials.[27,30-33] The challenges of assessing ENMs with existing test methods has also been recognized by the test guideline program of the Organization for Economic Co-operation and Development (OECD). The OECD launched the testing program of ENMs in 2007 to ensure that the approaches for hazard, exposure, and risk assessment for ENMs are science based and internationally harmonized.[34] One objective of this program is to explore the potential application of alternative methods for testing of ENMs. Validated methods have been developed for dissolved chemical substances, yet additional work is needed in some cases to adapt these methods for use with ENMs given their different behaviors in the assay system. For example, there is a possibility for ENMs to cause artifacts in these assays such as by adsorbing reagents of the test system,[35-38] damaging biomolecules after a cell assay has concluded (e.g., ENM induced photoactivity as a result of laboratory lighting),[21,39] or producing a signal (e.g., fluorescence or absorbance) similar to the measurand for the assay.[19,20] Biases caused by these artifacts and other important, yet often uncontrolled, aspects of the assay such as cell-seeding density may be the source of contradictory results on the toxicity of various ENMs. Another complication for the development of robust assays for ENMs is the dispersion protocol, which, if not standardized, may lead to different ENM suspensions even when using the same starting material.[17,18,40-42] Testing the ENMs with same composition but from different manufacturers may generate different results, hindering comparison among *in vitro* assay results even within a single laboratory.[43]

One approach that we have used previously to comprehensively evaluate potential sources of uncertainty for the 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-

tetrazolium (MTS) cell viability assay is cause-and-effect (C&E) analysis.[15] C&E analysis is an approach, originally utilized in quality manufacturing and propagation of measurement error in analytical chemistry.[44] It highlights through C&E diagrams, prepared based on expert judgment, aspects of the assay anticipated to most strongly influence the variability of the assay result. C&E diagrams, in particular, provide a graphical representation of potential sources of variability in assays. These diagrams can then be used to add process control measurements to an assay protocol to track key potential sources of variability.[15] Understanding these sources of variability and the robustness of the assay to minor, unintended changes in the protocol (e.g., temperature, stability of reagents, time range for incubation) enables scientifically informed, instead of *ad hoc*, choices about components of the assay to monitor, and allowable ranges for different steps of the assay (e.g., temperature or incubation duration). This supports reducing interlaboratory variability during interlaboratory testing and eventual approval for usage in regulatory testing.[45] For example, during sensitivity testing of the MTS assay, we determined that the cell pipetting caused the highest variability in the assay result.[15] These findings and the use of process control measurements supported successful interlaboratory testing of the protocol where outlier results from one laboratory could be traced to a different interpretation of a single assay step and corrected by revising the procedure.[46] However, it was unclear from this case study about how to apply this approach to other nanocytotoxicity assays and if the same control measurements would be sufficient to capture the most important steps in these assays, particularly if the assays required more complex steps than the relatively straightforward MTS assay.

To address this topic, a workshop took place in St. Gallen, Switzerland, during June 2015 entitled "Cause-and-effect analysis: A new approach for developing robust nanobioassays" with 17 participants from 4 countries. During the workshop, the application of C&E analysis was evaluated on four additional cell-based assays, selected by the workshop participants, for use with ENMs: the DCF (2',7'-dichlorofluorescein) assay, an enzyme-linked immunosorbent assay (ELISA) for measuring interleukin-8 (IL-8), a flow cytometry assay (Annexin V/Propidium Iodide (PI)), and the Comet assay. These assays can be used to measure a possible cascade of events that can occur in cells impacted by ENMs through oxidative stress, a condition in which the amount of intracellular ROS produced overwhelms the cells' antioxidant defense capacities.[47,48] Oxidative stress has been shown in numerous nanotoxicity studies to be the principal mechanism causing toxicity given the potential for a broad range of ENMs to produce reactive oxygen species (ROS) or activate redox reactions. Because of this, reactivity is among the most promising intrinsic ENM properties used in ENM hazard categorization.[47,49-52] Given that it is not yet clear which combination of assays measuring oxidative stress is most predictive to estimate *in vivo* effects,[47] the previously mentioned four assays were evaluated: the DCF assay which directly measures intracellular ROS, the ELISA which measures the increase in cytokines indicative of inflammation, the flow cytometry assay which measures the mechanism of cell death which can be caused if the cells are exposed to sustained oxidative stress, and the Comet assay which measures DNA damage, another potential outcome of increased intracellular ROS levels. However, it should be noted that positive responses with these assays can also be caused by mechanisms other than oxidative stress. For each assay, we performed a C&E analysis and designed control measurements to assess for ENM interference with the assay. The extent to which data from positive chemical control samples can provide insights into how well the assay is functioning is also discussed. Lastly, next steps are described that could be taken with these assays to more fully evaluate different expected sources of variability.

Overview of commonly used in vitro-based cytotoxicity assays

An overview for each of the four assays investigated during the workshop is provided in Table 1 and each assay will be briefly described. The full protocols for each assay are provided in the Supporting Information (SI). Flow charts which describe each step of the assay are shown in Figure 1. More detailed descriptions and protocols for each of these assays are available in the SI. The DCF assay is designed to measure production of reactive oxygen species (ROS) by ENMs, one of the most common mechanisms of ENM toxicity. The assay works by assessing the rate at which a non-fluorescent dye is chemically reduced by ROS to form a fluorescent dye. This assay can be conducted under acellular or cellular conditions. In the acellular version of the assay, there are additional steps to remove the diacetate moiety while deacetylation occurs intracellularly in the cellular version of the assay. In the ELISA, the release of cytokines, molecules that are indicative of inflammatory reactions, is quantified. The production of a specific cytokine, in this example interleukin-8 (IL-8), is measured using the sandwich ELISA method which works as follows: 1) adsorbing a primary antibody onto the surface of a high-affinity binding microwell plate, 2) having the antibody bind the protein of interest in the cell culture supernatant, 3) adding a second antibody to bind to the same protein of interest but at a different epitope, 4) adding horseradish peroxidase linked to avidin to initiate an enzymatic reaction with tetramethylbenzidine (TMB) as the substrate, and 5) quantifying the TMB using absorbance on a plate reader. In the third assay, flow cytometry can be used to evaluate the quantity of apoptotic and necrotic cells. Cells are treated with markers to distinguish between (i) viable cells, (ii) cells which are not viable but have the membrane intact (apoptotic cells), and (iii) cells which are not viable and also have undergone membrane disintegration (necrotic cells). The cells are subsequently analyzed using flow cytometry. In contrast to the *in vivo* situation, apoptotic cells cannot be removed by tissue macrophages *in vitro.* Their membranes start to disintegrate and cells stain double-positive for both markers (late apoptotic cells). Lastly, the purpose of the Comet assay is to assess the potential genotoxicity of a compound through measuring the degree of DNA damage that has occurred. The DNA integrity of cells is evaluated after gel electrophoresis, and the length and quantity of DNA in the Comet tail relates to the extent of DNA damage.

Control charting data

Control charting data of the positive and negative control results (i.e., results for these control measurements from assays performed on multiple different days) from the four assays obtained at Empa using A549 cells are provided in Figures S1 through S4. The control charting data was also evaluated to assess if there was a statistically significant correlation among the control charting data for different components of data in the control charts (e.g., the negative control data and the positive control data) (e.g., Figures S5 and S6). The nonparametric Spearman's rank test (GraphPad Prism) was applied to the data to determine if there was a trend in the rank of data, and in some cases a linear regression was used. Discussion of these results is provided in the SI.

Overall, these results indicate that all of these assays could benefit from decreasing their variability and increasing the within-laboratory reproducibility. Therefore, the assays were evaluated in greater detail in subsequent sections to discuss possible options for improving their reproducibility and for enabling their use with ENMs.

Evaluation of sources of variability in the assays

Only two steps are similar among all of the assays: the initial cell seeding step and the step related to treating the cells with the ENM and positive chemical controls (see Figure 1), although the ENMs are dispersed in different media among the assays prior to cell treatment. To assess the ENM dispersion, it is important to characterize, preferably using orthogonal techniques, the suspended ENMs such as the extent to which they have agglomerated and their interactions with serum proteins (if in the media). It is important to note though that these measurements are challenging and the results may vary among analytical techniques. Overall, improvements in the repeatability for the ENM dispersion, cell seeding, and dosing steps could help improve the precision of all of the assays. However, the extent that the variability would be decreased would depend on the relative contribution of these steps to the total variability for each assay. It is also evident when viewing the flow charts that there are substantial differences in the number of steps in the assays with the Comet assay and the ELISA method containing the highest number. The instrument used to obtain the final assay readout also varies with two of the assays requiring absorbance measurements, while the other assays utilize flow cytometry, fluorescence or microscopic analysis.

The C&E diagrams showed similar sources of variability among assays and key differences (Figures 2 through 6). The aspects of the branch colored in orange are designed to highlight differences compared to the previously published C&E diagram for using the MTS assay with ENMs (Figure 2). For all of the assays except for the MTS assay, the positive chemical control to be used to elicit a similar response as the potential impact of the ENM was a significant source of variability. Therefore, this topic will be discussed in depth in subsequent sections.

These C&E diagrams were created based upon the specific protocols available in the SI. While there are factors such as the media used or the serum percentage that vary among protocols, these factors are not discussed in depth here since these experimental details are specified in the protocols. Similarly, the use of high throughput testing may help minimize variability in assay results comparing different ENMs if the suspensions are tested at the same time,[53] but that is not possible using the protocols described here. In addition, the branches for the C&E diagrams were determined by expert judgement and can be refined based upon robustness testing of the assays[54] since some of the expected sources of variability may not have a substantial contribution. These sources of variability focus on unintended deviations from the assay protocol (e.g., random variability from pipetting cells among multiple pipetting ejections, or misinterpreting a step of the protocol such as cell rinsing during the MTS assay[46]). The C&E diagrams do not cover mistakes performing the assay protocol (e.g., pipetting double the intended volume or if the cells are contaminated such as with mycoplasma), which should be avoided through a laboratory's quality control system. In other words, to the best of his or her knowledge, the operator executes the protocol correctly, but nevertheless, unnoticed or unavoidable sources of variability lead to varying assay results.

To address the sources of uncertainty revealed during the C&E analysis, sensitivity testing can be conducted to quantify the amount of variability from these different sources and control experiments can be performed to evaluate potential artifacts from ENM testing. Overall, potential artifacts from ENM testing depend upon the individual steps in an assay and also the analytical instrument (e.g., plate reader) used to perform the measurement. Thus, there are not overarching control experiments that would be relevant for all four assays. Nevertheless, the conceptual process for identifying control experiments related to artifacts from ENM exposure for these four assays is similar with three primary questions: 1) Is the ENM able to process or modify probe molecule in the absence of cells? 2) Is the ENM fluorescent or absorbent by itself? 3) Is the ENM reducing or increasing an existing signal? These considerations and insights from the C&E diagrams were used to delineate control measurements for each assay several of which are shown in Tables 2 through 5 and are discussed for each assay in the following sections.

It is valuable to note that there are key control measurements for instrument performance specific to the instruments used for each of the assays. For example, it is important to ensure that the plate reader is calibrated and to confirm that it provides a linear response for the conditions being used in the assay (e.g., for the fluorescent DCF molecule for the DCF assay). For older plate readers, there is typically only a limited number of settings available, and thus the specific suggested excitation and emission wavelengths for a dye may not be available. Therefore, it is often necessary to determine the optimal settings for each system. It should be noted though that differences in the excitation/emission settings for plate readers can be a source of increased variability among laboratories. There are similar considerations that are relevant for a fluorescent microscope (e.g., linearity, benchmarking) and the flow cytometer (e.g., potential for spillover for the dyes used and assessment of the size of the cells to be tested by measuring their side and forward scattering to determine the binning). Since these aspects of instrument maintenance and calibration are not specific to these assays, they will not be discussed in additional detail in the subsequent sections.

In the following sections, each assay will be separately discussed in depth including a general discussion of results from the C&E analysis, control measurements identified from the C&E analysis for the assay (including ENM-specific measurements), and strategies to improve the quality of each assay. It is also important to note that some of these control measurements can be conducted in preliminary experiments such as to identify optimal instrument settings to refine the protocol or to assess for the potential of ENM-induced artifacts. If such artifacts are observed and the impact is substantial, this may indicate that modifications to the protocol are needed or that the assay cannot be used with this ENM. In other cases, modest ENM-induced artifacts may be corrected for using process control measurements, measurements that are made each time the assay is conducted such as the performance of the positive chemical control to provide evidence that the assay is functioning as expected. For example, it may be possible to correct for bias from ENM settling during the MTS assay by performing background subtraction using wells dosed with the same ENM concentration but without cells which undergo all subsequent assay steps.[46]

Detection of reactive oxygen species by H_2DCF-DA assay (DCF assay)

For the DCF assay, a main source of variability relates to the chemical reactivity of the independent positive chemical control (Figure 3). There are no known positive chemical controls for generating consistent quantities of ROS species in this assay, thus hindering the comparability of results with the DCF assay among laboratories or across time in a single laboratory.[14] What is needed is a control substance that can produce a consistent amount of ROS within the cells, is readily available, not cytotoxic, and can be quantified in solution.[14] A positive chemical control that fulfills these criteria could be used as a calibrator in a concentration series to allow the comparability of results across time and space by comparing a response from this positive chemical control such as an EC_{50} value. Given the inability to find a positive chemical control that can yield a consistent amount of ROS in this assay, the results from this assay are not quantitative and cannot be readily compared across experiments or laboratories. Other unique sources of variability in this assay relate to the instability of the H_2DCF-DA reagent with time and as a result of its potential to be degraded by laboratory light (unpublished results). This can result in variability in the assay results that are challenging to quantify since a positive chemical control capable of generating a consistent quantity of ROS species is not available.

There are several key control measurements for the DCF assay in addition to positive chemical control measurements and assessments of potential biases or artifacts from testing ENMs. These control measurements can be assessed during preliminary experiments to assess the robustness of the assay. It is valuable to assess whether loading the dye in the light or dark would influence the assay results, although the H_2DCF-DA molecule, which is first added to the assay, is more stable than the highly light-sensitive fluorescent DCF molecule, which is measured by the assay. In addition, it is also possible to compare the results among H_2DCF-DA provided by different manufacturers to assess to what extent the source of this assay reagent impacts the results. Another consideration is the potential impact of repeated freeze/thaw cycles on the H_2DCF-DA molecules. For each of these topics, it is necessary to keep other factors (e.g., cell handling) as constant as possible or to use cell-free control measurements to assess changes in the H_2DCF-DA molecule. For example, if repeated freeze-thaw cycles are to be tested, one needs to make sure that apparent differences in the H_2DCF-DA with time are not due to changes in the chemical (e.g., Sin-1) that was used to induce ROS which may itself change during storage. As a result of the high Sin-1 reactivity, it is possible that the pipetting time needed to aliquot Sin-1 throughout a whole 96-well plate could be sufficiently long to result in changes from the first to the last wells to which Sin-1 was added. This can hinder determining if changes in the signal measured are from the Sin-1 or changes to the H_2DCF-DA molecule.

There are also several key ENM-specific control experiments for the DCF assay that should be evaluated during preliminary experiments. It is valuable to assess the time needed for ENMs to contact the cells. While the H_2DCF-DA molecule is first allowed to enter the cell after which point the cells cleave the DA and then the excess dye is washed away prior to ENM addition, the time points after the ENM addition should allow sufficient time for the ENM to have contacted the cell before measuring the assay output. It is possible though that the ENMs can produce ROS and impact the deacetylated H_2DCF in the absence of cells, thereby causing a false positive signal. This can be assessed by performing an acellular experiment with the dispersed ENMs and the deacetylated H_2DCF molecule. A positive signal in this acellular control experiment indicates that the ENM can produce ROS extracellularly or that the reactive

surface of the ENM can process the dye although not necessarily through the production of ROS. If a positive result in this acellular assay is obtained, a positive response in the cellular assay should be interpreted with caution and additional measurements such as antioxidant biomarkers (e.g., vitamin C or N-acetyl-cysteine) should also be performed before concluding that the ENM causes intracellular ROS. It is also important to assess the extent to which ENMs in the absence of cells and assay reagents can produce results similar to the DCF signal (namely excitation at a wavelength of 485 nm and emission at a wavelength of 528 nm) or can quench the existing fluorescent signal from the DCF molecule [55,56]. These potential biases can be tested by measuring the fluorescent signal for the dispersed ENMs by themselves or by incubating the fluorescent DCF molecule with dispersed ENMs at a range of ENM concentrations and assessing if there is a change in the fluorescent signal.

Numerous steps can be undertaken to further improve the quality of the DCF assay results such as comparing results obtained using reagents from different manufacturers, assessing the impact of freeze-thaw cycles on the H_2DCF-DA, and conducting the ENM related control experiments described above. These ENM related control experiments are critical for proving that the results have not been biased. However, there are limitations regarding the extent to which this assay can be improved given its challenges related to the instability of positive chemical controls which induce ROS thus hindering the ability of results from this assay for different experiments to be compared.[14] Given the sources of uncertainty in this assay, it will be challenging to comprehensively determine that an ENM does not produce ROS above a threshold since the results are currently not comparable among measurements. For example, if ROS is not detected and it is determined that this is not an artifactual result, this finding could stem from multiple factors: insufficient ENM contact with the cells at the time point that ROS is measured given that an increasing fraction of the delivered ENM dose contacts with the cell at longer time points, a lack of ROS production at the time the assay is conducted if the cells have increased their antioxidant to counteract ROS generated by contact with ENMs, the increase in intracellular ROS being too low to detect, or that the ENM does not have the capacity to produce ROS. The highest possible quality for this assay is for there to be grouping of the ROS generating potential perhaps into no detectable ROS generation, weak ROS generation, and strong ROS generation.

In addition, this assay is typically conducted in the absence of serum, because the serum can contain esterases which can cleave DA from H_2DCF-DA leaving the molecule vulnerable to extracellular ROS attack thereby producing a false positive signal.[14,57] Therefore, the results may be challenging to compare to other assays which do use serum because the serum could change the surface of the ENM by producing a protein corona which can impact the ENM interactions with cells, agglomeration behaviors, and *in vitro* dosimetry.[46,58,59] Nevertheless, it may be possible though to rank the ROS generation of ENMs within a single laboratory if the ENMs are tested during the same experiment or if the results are shown to be comparable across multiple experiments. This assay can also be conducted in two distinct approaches: load the cells with H_2DCF-DA and then apply the ENMs or apply the ENM and then later add the dye. These two approaches can be used to analyze different types of ROS production. If the dye is added first, it is possible to test for an increase in ROS after the first ENM internalization or contact. If the ENM is added first, the initial increase in ROS is not detected, but a longer-term increase in ROS can be measured. However, both of these approaches are very sensitive to kinetics, because the time that the

initial increase in ROS occurs depends upon the rate at which the ENMs contact the cells, while longer-term ROS can be mitigated through cellular antioxidant mechanisms. Overall, a positive response from this assay does provide a general indication that ROS are being produced within cells which can be evaluated in more depth if needed with other assays such as more complicated and expensive techniques for acellular measurements (e.g., electron paramagnetic resonance).

Detection of the production and excretion of the pro-inflammatory cytokine IL-8 by ELISA

The ELISA method has several distinct differences from the other assays (Figure 4). For example, it is challenging to assess the purity of the TNF-α positive chemical control. This recombinant protein initiates a cell signaling cascade that results in the cellular production and excretion of IL-8. However, it is not straightforward to assess the purity of TNF-α, and the reagent quality may differ among manufacturers (additional discussion of the positive chemical control is provided in the SI). In addition, other factors such as the quality of the high binding plates can also influence assay results. Similarly, there is substantial variability in the quality of the antibodies used in the assay, but there is not a single universally accepted method for assessing antibody quality.[60] Antibodies are complex biomolecules that are challenging to fully characterize. Thus, differences among antibodies produced by different manufacturers could influence the assay results. Given that the ELISA protocol requires 22 washing steps, the variability caused by each washing step has the potential to be additive and substantially increase the variability of the assay results. The impact of manual versus automated washing on the variability of assay results can be evaluated during preliminary experiments.

One key consideration for this assay is whether to use a commercial kit directly as specified or to titrate antibodies to try to achieve a better sensitivity. For both approaches, it is important to perform a dose-response curve, often known as a standard curve, to evaluate the signal obtained in wells with a known quantity of the cytokine added. The standard curve is needed to assess the performance of the ELISA procedure during that specific experiment (e.g., is the response linear and is the sensitivity sufficient) and to quantify the protein mass in the supernatant of the control and ENM exposed cells. The information from the standard curve can also be used for control charting to evaluate the assay performance across time. It is important to add that the substrate used in the assay (e.g., TMB: tetramethylbenzidine; OPD: o-Phenylenediamine dihydrochloride; or ABTS: Azinobis-ethylbenzothiazoline-6-sulfonic acid-diammonium salt for HRP activation) could result in different sensitivities for the same kit. Several other factors can also impact the assay results including the time allowed for the enzyme-substrate reaction, whether the enzymatic reaction is chemically stopped and the HRP activity (which can change among lots and suppliers and with storage time).

Comparable to the DCF assay, it is possible that the ENMs could give a signal similar to the measurand of interest (e.g., absorbance at 630 nm for the measurement of IL-8) or that the ENMs could interact with the assay reagents and decrease the signal strength. This can be evaluated during preliminary experiments for each ENM to be tested. It is possible to conduct acellular control experiments to assess the extent to which ENMs at the cell media concentrations used in the assay provide an absorbance signal at 630 nm. In addition, it is possible to expose the cells to the positive chemical control (typically LPS) for a set exposure duration, remove the supernatant, add ENMs to the supernatant, and

then perform the ELISA procedure to assess if the presence of ENMs could cause a bias in the ELISA procedure. Unlike for the DCF assay, another potential bias is for the ENMs to adsorb the IL-8 produced by the cells during the course of the cell exposure, thereby decreasing the measured IL-8 concentration.[55] This can be evaluated by conducting adsorption experiments in the cell media with ENMs at the highest concentration of ENMs to be tested and with a known concentration or range of concentrations for the IL-8 cytokine and then performing the ELISA assay to assess the IL-8 recovery.

For ELISA, there have been varying degrees of agreement among the laboratories in the two interlaboratory comparison studies conducted with ENMs.[24,61] One study showed that all seven laboratories observed significantly increased cytokine levels after cellular exposure to the highest concentration of TiO_2 ENMs as compared to the control samples but the cytokine concentration after TiO_2 ENM exposure varied by roughly an order of magnitude.[24] In addition, these laboratories showed inconsistent findings about the capacity for certain types of multiwall carbon nanotubes to increase the cytokine concentration. In a second interlaboratory comparison, there were mixed results with two laboratories showing a significantly increased cytokine concentration and two other laboratories showing no effect for cells exposed to silver ENMs.[61] It was unclear why the absolute cytokine concentration varied substantially among laboratories or why there was not better interlaboratory agreement for the second study. However, this variability among laboratories in the absolute cytokine concentration is similar to the intralaboratory variability shown in Figure S2. These findings suggest that the ELISA can potentially yield qualitative interlaboratory agreement (e.g., “yes” or “no” for an increased cytokine concentration compared to the positive chemical control), but it is unclear if this assay can be refined to yield a quantitative agreement. Additional experiments to more carefully evaluate the impact of different sources of variability could be used to further refine the assay and reveal process control measurements to include in the ELISA method to quantify sources of variability more thoroughly during each experiment. However, it is currently unclear which combination of process control measurements would help improve the comparability of results across experiments within a laboratory and among laboratories. It is possible that improving the experimental protocol to decrease the variability in the negative control data could lead to improved comparability, but given the lack of a correlation between the negative and positive control values, it is unclear to what degree this would improve the reproducibility of the assay. Evaluating the robustness of the dose-response relationship for the positive chemical control to changes in the assay protocol may be more valuable.

Detection of cell death mechanism by an Annexin V/PI flow cytometry assay

Unlike the ELISA and DCF assay, there are substantially more sources of significant uncertainty in the “instrument performance” branch, largely as a result of challenges related to achieving interlaboratory comparability among flow cytometry results. These challenges are not unique to ENMs. It is challenging with flow cytometry to develop benchmarks to enable comparability among instruments for multiple reasons: it is challenging to develop an absolute calibration; it is challenging to assess the instrument’s linear range of operation; background debris, which is produced by dead cells as they are degraded, may influence the results especially for assays with ENMs, which may have optical properties similar to the

dyes used to stain the cells; and ENMs may have unexpected interactions with the cell debris causing measurements potentially similar to apoptotic or necrotic cells. Multiple components of the assay protocol branch may also contribute to the overall uncertainty in the assay results such as the procedure used for preparation of the cells for the flow cytometry analysis including cell harvesting and resuspending the cell pellet, and the quality of the PI and Annexin V reagents given that the Annexin V is only stable in a special buffer. One key step is to determine the gating strategy to be used for all experiments that will be directly compared. The gating strategy and compensation procedures will be especially important for comparisons of results among laboratories. Determining a standardized strategy for gating and compensation is a topic of ongoing research.[62,63]

ENMs may impact specific fluorophores preferentially resulting in either enhanced (false positive) or quenched (false negative) fluorescence signals. For example, Annexin V can be obtained with different fluorophores such as FITC or PE, and "simply" changing the fluorophore may prevent interference reactions. Similar observations have been made for the DNA intercalating dye PI. Changing the "necrosis specific" dye to 7-aminoactinomycin D (7-AAD) eliminated the interference reaction of SiO_2 particles with PI.[64] The potential for interference of the ENM on the dyes can be evaluated during preliminary experiments. Furthermore, the overall viability status of the cells used for flow cytometry should be assessed by a second independent method such as the trypan blue exclusion assay. This can be conducted each time the assay is performed if feasible. A massive PI signal in flow cytometry without any trypan blue positive cells under the microscope within the same sample indicates ENM interference in at least one of the assays, thereby necessitating further investigations.

One important control experiment for this assay is to evaluate if the cell harvesting procedure may produce debris that could then interact with ENMs and hinder the quantification of the apoptotic, or the late apoptotic and necrotic cells. This could be evaluated by a 0 h control experiment in which the cells are immediately harvested after ENM addition and the assay performed. It is possible that ENM agglomeration, ENM interactions with cell debris, or ENM interactions with stained or unstained cells could impact gating choices and subsequent analyses.[55]

There have not been any interlaboratory comparisons conducted with flow cytometry to assess necrotic and apoptotic cells after ENM exposure to our knowledge. Nevertheless, several potential topics to refine this assay have been discussed in this manuscript, all of which could support the successful conduct of an interlaboratory comparison. Some of the topics described would be relevant for a range of flow cytometry measurements (e.g., gating), while most of the control experiments relate to ENM-relevant issues such as interference between the ENM signal and that of the measurand. In addition, it is possible to spike in ENMs to the cells obtained during control experiments prior to flow cytometry analysis to evaluate if the presence of ENMs enhances or quenches the signal for apoptotic or necrotic cells.[64]

Detection of cellular DNA damage by the Comet assay

There are numerous components in the “Assay protocol” branch (Branch 5) of the Comet assay C&E diagram that can contribute to uncertainty in assay results (see Figure 6). While some of these are similar to sources of variability determined for other assays (e.g., the cell harvesting step is shared with

429 the flow cytometry assay), many of them relate to the unique aspects of this assay such as the alkaline
430 cell lysis step in the protocol. In addition, the instrument used to perform the gel electrophoresis differs
431 from the instruments used in the other assays. Similar to the other assays, the effectiveness of the positive
432 chemical control is also a major source of variability for the Comet assay.

433 There are numerous control measurements that can be performed to assess the performance of
434 specific aspects of the Comet assay each time the assay is run. There are some commercially available
435 control cells with variable amounts of DNA damage (e.g., leukocytes treated with etoposide[65] from
436 Trevigen). While these controls are different from positive chemical controls conducted by exposing cells
437 to chemicals during the assay itself, they can provide information about the consistency of the
438 performance of other steps in the assay protocol (e.g., the electrophoresis step). In addition, different
439 dyes to stain DNA after electrophoresis can be utilized. It is important to choose a dye that intercalates
440 into double- and single-stranded DNA and to maintain consistency among the dyes within a study since
441 different dyes may yield varying results.

442 The equipment used to perform the gel electrophoresis can also impact the assay results. For
443 example, it is possible that there is a change in temperature during this step of the assay or that there is
444 heterogeneity in the voltage across the gel. Control measurements could be conducted to evaluate the
445 heterogeneity in the voltage in the gel electrophoresis step by placing slides of control samples at different
446 locations in the tank to see if similar results are achieved. It may also be possible to take steps to control
447 the temperature and pH of the test setup or to measure them before and after the assay is performed to
448 assess if there was a change. These measurements can be performed during preliminary experiments. It
449 may also be helpful to measure the temperature during the gel electrophoresis step each time the
450 experiment is performed to ensure it is consistent among experiments.

451 Another key set of control measurements relates to the microscopic analysis of the comets.[66] For
452 example, microscopic settings such as the focus and camera exposure time can cause variations in
453 percentage DNA in the tail by up to 40 % although there are steps that can be taken to improve the
454 reproducibility of this component of the assay.[66] Selection of the comets can also impact the results
455 especially if the slides are not scored blindly (i.e., without knowing which sample the operator is scoring).
456 One approach to minimize this source of variability is to automate the comet selection. It is also possible
457 to analyze the same sample multiple times and potentially using multiple approaches (e.g., manually or
458 with automation, or with using different Comet selection parameters for the automated Comet selection)
459 to assess the variability of this step.

460 There are several relevant potential mechanisms through which ENMs may cause artifacts or
461 biases in the Comet assay: 1) the presence of ENMs may influence the DNA migration rate during the gel
462 electrophoresis step, 2) ENMs may associate with the nucleus, migrate themselves during the gel
463 electrophoresis step, and be misinterpreted as damaged DNA or DNA in the Comet head during the
464 microscopic analysis, and 3) reactive ENMs may damage DNA during the processing steps after the cell
465 exposure (for example through photoactive ENMs being activated by laboratory light during the cell
466 processing steps and causing DNA damage) and then this damage is misinterpreted as having occurred
467 during the cell exposure.[21,68,69] There are several control measurements that could be performed to

evaluate the extent to which these biases or artifacts have occurred. First, a 0-h control experiment could be performed in which ENMs are added to the cells followed by immediate processing of the cells.[70,71] If there is an increase in apparent DNA damage for the cells with ENMs as compared to untreated cells, this indicates that the ENMs have caused an artifact. It may be possible to wash the nuclei after the treatment process to remove ENMs, but this approach has only recently been tested.[72] Second, the Comet tail could be analyzed microscopically such as with hyperspectral imaging analysis to evaluate the extent to which ENMs are present in the tail (e.g., using hyperspectral imaging analysis) and the extent to which they could cause apparent DNA damage.[73] Third, for potentially photoactive ENMs, it is possible to treat a large number of replicates and then conduct the cell processing for some samples under laboratory light conditions and for other samples with lighting that has a narrow spectrum or reduced light intensity designed to minimize photoactivation.[21,68,74]

One of the main approaches that could be used to improve the quality of the Comet assay is the use of cells with carefully controlled DNA damage such as through irradiation to investigate the impact of different aspects of the assay. The intrinsic variability of biological samples (i.e., cells within a single replicate) typically yields a broad dispersion of results for each replicate. Sample-to-sample variability may also be substantial. Therefore, a batch of frozen, single-use aliquots of consistent cell samples ("reference cells") could, for example, be used to evaluate the impact of different factors with regards to the gel electrophoresis step such as the impact of voltage, temperature, time for unwinding and electrophoresis, and low melting agarose concentration. It could further be used to improve the automation and reliability of the microscopy steps which could produce similar results regardless of the sample operator.

A range of potential biases or artifacts have been identified in the Comet assay when used with testing ENMs, but these control experiments and how to handle artifacts when observed can be better elucidated. The extent to which these can be mitigated such as through washing the DNA after the exposure prior to the gel electrophoresis step should be further examined. It is also unclear if it is possible to correct for biases from the ENMs if they are observed or if an alternative genotoxicity assay should be utilized instead. Additional analysis of Comet assay slides with ENMs in the tails could be evaluated as described above to assess the impact of ENMs in the Comet assay. In interlaboratory results from the Comet assay with ENM exposure, there have been mixed results with some laboratories showing non-comparable results.[75]

Conclusions

The steps described in this paper outline a process of evaluating the quality of *in vitro* toxicological assays for use with ENMs using C&E diagrams and describe specific measurements that can be taken to improve the quality of each of the four assays. Producing a C&E diagram is a key step that can be utilized in the development of any robust toxicological assay. The C&E diagram can then guide subsequent robustness testing to quantify different sources of uncertainty. Based on this information, it is possible to design a protocol that includes preliminary control experiments which may need to be performed with each ENM to, for example, identify if potential artifacts are observed, and also process control measurements which yield information about the assay performance each time it is performed. It is important to note that many of the branches in the C&E diagrams are similar among the different assays

and were the same as for the MTS assay.[15] Therefore, it may be easier to prepare C&E diagrams for additional assays once it has been conducted for one. In addition, improvements in the precision for shared steps among the assays can result in improvements for all of the assays.

The degree to which the precision and robustness of the assay for use with ENMs needs to be improved for a specific application depends upon ensuring that the assay is fit for purpose. For example, assays used for screening purposes may require less precise output and qualitative answers (e.g., yes or no) may be sufficient. For use in quantitative risk assessment or for replacement of *in vivo* assays, the quality of the assay may need to be substantially higher. A promising overall strategy is for some combination of these assays to be used in an integrated approach for testing and assessment (IATA) for use in screening ENMs for additional testing using *in vivo* assays (e.g., for inhalation exposure) or to predict results from *in vivo* assays. It is unlikely that any individual assay will be able to fulfill either of these purposes, but combinations of assays for use in IATAs are more promising. Therefore, it is challenging to predict *a priori* which of these assays will be most helpful for these purposes in the absence of test results and comparisons to a specific set of *in vivo* results.

In general, the use of orthogonal methods such as testing cytotoxicity with different approaches is highly valuable for nanocytotoxicity studies and can help identify artifactual results. Nevertheless, increasing the number of assays does also increase the cost and resources required for each ENM to be tested. In addition, it is impossible given the broad range of ENMs that can be synthesized and their variable behaviors and properties to provide descriptive information about which potential biases or artifacts will be the most important for each assay. Previously published results for ENMs with similar properties will likely give guidance about the likelihood of a certain ENM to cause a certain artifact but control experiments need to be conducted on a case-by-case basis. Another key topic for developing assays is their statistical robustness and design to minimize the frequency of false positive and false negative results. Different statistical methods can be utilized for qualitative (e.g., high, low, or no effect) as compared to quantitative results (e.g., an EC_{50} value with a defined uncertainty) to ensure the assay's statistical robustness.[76,77] The approach described in this paper can also inform the evaluation of other biological assays for use with ENMs and support their refinement and eventual validation.

Supporting information

Materials and methods section and detailed protocols for all assays, supporting results related to evaluation of the control charting data, supplemental discussion related to positive chemical control considerations, control charting data for all four assays, summary data for different Sin-1 concentrations for the DCF assay, and analysis of control charting data for the ELISA method. This information is available free of charge via the Internet at http://pubs.acs.org.

Acknowledgements

We acknowledge funding from the NanoScreen Materials Challenge co-funded by the Competence Centre for Materials Science and Technology (CCMX).

Certain commercial products or equipment are described in this paper in order to specify adequately the experimental procedure. In no case does such identification imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that it is necessarily the best available for the purpose.

Conflict of interest statement

The authors declare no conflicts of interest.

Corresponding Author: Elijah J. Petersen, E-mail: elijah.petersen@nist.gov, Phone: 301-975-8142

References

1. Pavlidou, S.; Papaspyrides, C. D., A review on polymer-layered silicate nanocomposites. *Prog. Polymer Sci.* **2008,** *33* (12), 1119-1198.
2. Petersen, E. J.; Pinto, R. A.; Shi, X. Y.; Huang, Q. G., Impact of size and sorption on degradation of trichloroethylene and polychlorinated biphenyls by nano-scale zerovalent iron. *J. Hazard. Mater.* **2012,** *243*, 73-79.
3. Petersen, E. J.; Lam, T.; Gorham, J. M.; Scott, K. C.; Long, C. J.; Stanley, D.; Sharma, R.; Liddle, J. A.; Pellegrin, B.; Nguyen, T., Methods to assess the impact of UV irradiation on the surface chemistry and structure of multiwall carbon nanotube epoxy nanocomposites. *Carbon* **2014,** *69*, 194-205.
4. Potts, J. R.; Dreyer, D. R.; Bielawski, C. W.; Ruoff, R. S., Graphene-based polymer nanocomposites. *Polymer* **2011,** *52* (1), 5-25.
5. Zou, H.; Wu, S. S.; Shen, J., Polymer/silica nanocomposites: Preparation, characterization, properties, and applications. *Chem. Rev.* **2008,** *108* (9), 3893-3957.
6. Froggett, S. J.; Clancy, S. F.; Boverhof, D. R.; Canady, R. A., A review and perspective of existing research on the release of nanomaterials from solid nanocomposites. *Part. Fibre Toxicol.* **2014,** *11*.
7. Kaiser, J. P.; Roesslein, M.; Diener, L.; Wick, P., Human Health Risk of Ingested Nanoparticles That Are Added as Multifunctional Agents to Paints: an In Vitro Study. *Plos One* **2013,** *8* (12).
8. Kuhlbusch, T. A. J.; Wijnhoven, S. W. P.; Haase, A., Nanomaterial exposures for worker, consumer and the general public. *Nanoimpact* **2018,** *10*, 11-25.

9. Mitrano, D. M.; Motellier, S.; Clavaguera, S.; Nowack, B., Review of nanomaterial aging and transformations through the life cycle of nano-enhanced products. *Environ. Intl.* **2015,** *77*, 132-147.
10. Nowack, B.; David, R. M.; Fissan, H.; Morris, H.; Shatkin, J. A.; Stintz, M.; Zepp, R.; Brouwer, D., Potential release scenarios for carbon nanotubes used in composites. *Environ. Intl.* **2013,** *59*, 1-11.
11. G. C. Waissi-Leinonen, E. J. Petersen, K. Pakarinen, J. Akkanen, M. T. Leppanen and J. V. K. Kukkonen, Toxicity of fullerene (C60) to sediment-dwelling invertebrate Chironomus riparius larvae, *Environ. Toxicol. Chem.*, 2012, **31**, 2108-2116.
12. Pakarinen, K.; Petersen, E. J.; Alvila, L.; Waissi-Leinonen, G. C.; Akkanen, J.; Leppänen, M. T.; Kukkonen, J. V. K., A screening study on the fate of fullerenes (nC60) and their toxic implications in natural freshwaters. *Environ. Toxicol. Chem.* **2013,** *32* (6), 1224-1232.
13. B. C. Nelson, E. J. Petersen, B. J. Marquis, D. H. Atha, J. T. Elliott, D. Cleveland, S. S. Watson, I. H. Tseng, A. Dillon, M. Theodore and J. Jackman, NIST gold nanoparticle reference materials do not induce oxidative DNA damage, *Nanotoxicology*, 2013, **7**, 21-29.
14. Roesslein, M.; Hirsch, C.; Kaiser, J. P.; Krug, H. F.; Wick, P., Comparability of in Vitro Tests for Bioactive Nanoparticles: A Common Assay to Detect Reactive Oxygen Species as an Example. *Inter. J. Mol. Sci.* **2013,** *14* (12), 24320-24337.
15. Rösslein, M.; Elliott, J. T.; Salit, M.; Petersen, E. J.; Hirsch, C.; Krug, H. F.; Wick, P., Use of Cause-and-Effect Analysis to Design a High-Quality Nanocytotoxicology Assay. *Chem. Res. Toxicol.* **2015,** *28* (1), 21-30.
16. Burden, N.; Aschberger, K.; Chaudhry, Q.; Clift, M. J. D.; Fowler, P.; Johnston, H.; Landsiedel, R.; Rowland, J.; Stone, V.; Doak, S. H., Aligning nanotoxicology with the 3Rs: What is needed to realise the short, medium and long-term opportunities? *Regul. Toxicol. and Pharmacol.* **2017,** *91*, 257-266.
17. Guggenheim, E. J.; Milani, S.; Röttgermann, P. J. F.; Dusinska, M.; Saout, C.; Salvati, A.; Rädler, J. O.; Lynch, I., Refining in vitro models for nanomaterial exposure to cells and tissues. *NanoImpact* **2018,** *10*, 121-142.
18. Haase, A.; Lynch, I., Quality in nanosafety - Towards reliable nanomaterial safety assessment. *Nanoimpact* **2018,** *11*, 67-68.
19. Ong, K. J.; MacCormack, T. J.; Clark, R. J.; Ede, J. D.; Ortega, V. A.; Felix, L. C.; Dang, M. K. M.; Ma, G. B.; Fenniri, H.; Veinot, J. G. C.; Goss, G. G., Widespread Nanoparticle-Assay Interference: Implications for Nanotoxicity Testing. *Plos One* **2014,** *9* (3).
20. Petersen, E. J.; Henry, T. B.; Zhao, J.; MacCuspie, R. I.; Kirschling, T. L.; Dobrovolskaia, M. A.; Hackley, V.; Xing, B.; White, J. C., Identification and Avoidance of Potential Artifacts and Misinterpretations in Nanomaterial Ecotoxicity Measurements. *Environ. Sci. Technol.* **2014,** *48* (8), 4226-4246.
21. Petersen, E. J.; Reipa, V.; Watson, S. S.; Stanley, D. L.; Rabb, S. A.; Nelson, B. C., DNA Damaging Potential of Photoactivated P25 Titanium Dioxide Nanoparticles. *Chem. Res. Toxicol.* **2014,** *27* (10), 1877-1884.
22. Riebeling, C.; Piret, J.-P.; Trouiller, B.; Nelissen, I.; Saout, C.; Toussaint, O.; Haase, A., A guide to nanosafety testing: Considerations on cytotoxicity testing in different cell models. *Nanoimpact* **2018,** *10*, 1-10.
23. Wang, X.; Duch, M. C.; Mansukhani, N.; Ji, Z.; Liao, Y.-P.; Wang, M.; Zhang, H.; Sun, B.; Chang, C. H.; Li, R.; Lin, S.; Meng, H.; Xia, T.; Hersam, M. C.; Nel, A. E., Use of a Pro-Fibrogenic Mechanism-Based Predictive Toxicological Approach for Tiered Testing and Decision Analysis of Carbonaceous Nanomaterials. *ACS Nano* **2015,** *9* (3), 3032-3043.
24. Xia, T.; Hamilton, R. F.; Bonner, J. C.; Crandall, E. D.; Elder, A.; Fazlollahi, F.; Girtsman, T. A.; Kim, K.; Mitra, S.; Ntim, S. A.; Orr, G.; Tagmount, M.; Taylor, A. J.; Telesca, D.; Tolic, A.; Vulpe, C. D.; Walker, A. J.; Wang, X.; Witzmann, F. A.; Wu, N.; Xie, Y.; Zink, J. I.; Nel, A.; Holian, A., Interlaboratory Evaluation of

in Vitro Cytotoxicity and Inflammatory Responses to Engineered Nanomaterials: The NIEHS Nano GO Consortium. *Environ. Health Perspect.* **2013,** *121* (6), 683-690.
25. Collins, A. R.; Annangi, B.; Rubio, L.; Marcos, R.; Dorn, M.; Merker, C.; Estrela-Lopis, I.; Cimpan, M. R.; Ibrahim, M.; Cimpan, E.; Ostermann, M.; Sauter, A.; El Yamani, N.; Shaposhnikov, S.; Chevillard, S.; Paget, V.; Grall, R.; Delic, J.; Goni-de-Cerio, F.; Suarez-Merino, B.; Fessard, V.; Hogeveen, K. N.; Fjellsbo, L. M.; Pran, E. R.; Brzicova, T.; Topinka, J.; Silva, M. J.; Leite, P. E.; Ribeiro, A. R.; Granjeiro, J. M.; Grafstrom, R.; Prina-Mello, A.; Dusinska, M., High throughput toxicity screening and intracellular detection of nanomaterials. *Wiley Interdisc. Rev. Nanomed. Nanobiotechnol.* **2017,** *9* (1).
26. Harris, G.; Palosaari, T.; Magdolenova, Z.; Mennecozzi, M.; Gineste, J. M.; Saavedra, L.; Milcamps, A.; Huk, A.; Collins, A. R.; Dusinska, M.; Whelan, M., Iron oxide nanoparticle toxicity testing using high-throughput analysis and high-content imaging. *Nanotoxicology* **2015,** *9*, 87-94.
27. Nel, A. E.; Nasser, E.; Godwin, H.; Avery, D.; Bahadori, T.; Bergeson, L.; Beryt, E.; Bonner, J. C.; Boverhof, D.; Carter, J.; Castranova, V.; DeShazo, J. R.; Hussain, S. M.; Kane, A. B.; Klaessig, F.; Kuempel, E.; Lafranconi, M.; Landsiedel, R.; Malloy, T.; Miller, M. B.; Morris, J.; Moss, K.; Oberdorster, G.; Pinkerton, K.; Pleus, R. C.; Shatkin, J. A.; Thomas, R.; Tolaymat, T.; Wang, A.; Wong, J., A Multi-Stakeholder Perspective on the Use of Alternative Test Strategies for Nanomaterial Safety Assessment. *ACS Nano* **2013,** *7* (8), 6422-6433.
28. NRC, *Toxicity Testing in the 21st Century: A Vision and a Strategy*. The National Academies Press: 2007.
29. Godwin, H.; Nameth, C.; Avery, D.; Bergeson, L. L.; Bernard, D.; Beryt, E.; Boyes, W.; Brown, S.; Clippinger, A. J.; Cohen, Y.; Doa, M.; Hendren, C. O.; Holden, P.; Houck, K.; Kane, A. B.; Klaessig, F.; Kodas, T.; Landsiedel, R.; Lynch, I.; Malloy, T.; Miller, M. B.; Muller, J.; Oberdorster, G.; Petersen, E. J.; Pleus, R. C.; Sayre, P.; Stone, V.; Sullivan, K. M.; Tentschert, J.; Wallis, P.; Nel, A. E., Nanomaterial Categorization for Assessing Risk Potential To Facilitate Regulatory Decision-Making. *ACS Nano* **2015,** *9* (4), 3409-3417.
30. Clippinger, A. J.; Ahluwalia, A.; Allen, D.; Bonner, J. C.; Casey, W.; Castranova, V.; David, R. M.; Halappanavar, S.; Hotchkiss, J. A.; Jarabek, A. M.; Maier, M.; Polk, W.; Rothen-Rutishauser, B.; Sayes, C. M.; Sayre, P.; Sharma, M.; Stone, V., Expert consensus on an in vitro approach to assess pulmonary fibrogenic potential of aerosolized nanomaterials. *Arch. Toxicol.* **2016,** *90* (7), 1769-1783.
31. Landsiedel, R.; Kapp, M. D.; Schulz, M.; Wiench, K.; Oesch, F., Genotoxicity investigations on nanomaterials: Methods, preparation and characterization of test material, potential artifacts and limitations-Many questions, some answers. *Mut. Res.* **2009,** *681* (2-3), 241-258.
32. Landsiedel, R.; Sauer, U. G.; Ma-Hock, L.; Schnekenburger, J.; Wiemann, M., Pulmonary toxicity of nanomaterials: a critical comparison of published in vitro assays and in vivo inhalation or instillation studies. *Nanomedicine* **2014,** *9* (16), 2557-2585.
33. Krug, H. F.; Wick, P., Nanotoxicology: An Interdisciplinary Challenge. *Ang. Chemi. Intern. Edition* **2011,** *50* (6), 1260-1278.
34. OECD. **2018,** Testing Programme of Manufactured Nanomaterials, http://www.oecd.org/chemicalsafety/nanosafety/testing-programme-manufactured-nanomaterials.htm.
35. Belyanskaya, L.; Manser, P.; Spohn, P.; Bruinink, A.; Wick, P., The reliability and limits of the MTT reduction assay for carbon nanotubes-cell interaction. *Carbon* **2007,** *45* (13), 2643-2648.
36. Guo, L.; Bussche, A. V.; Buechner, M.; Yan, A. H.; Kane, A. B.; Hurt, R. H., Adsorption of essential micronutrients by carbon nanotubes and the implications for nanotoxicity testing. *Small* **2008,** *4* (6), 721-727.
37. Kroll, A.; Pillukat, M. H.; Hahn, D.; Schnekenburger, J., Interference of engineered nanoparticles with in vitro toxicity assays. *Arch. Toxicol.* **2012,** *86* (7), 1123-1136.

38. Worle-Knirsch, J. M.; Pulskamp, K.; Krug, H. F., Oops they did it again! Carbon nanotubes hoax scientists in viability assays. *Nano Lett.* **2006,** *6* (6), 1261-1268.
39. Shinohara, N.; Matsumoto, T.; Gamo, M.; Miyauchi, A.; Endo, S.; Yonezawa, Y.; Nakanishi, J., Is Lipid Peroxidation Induced by the Aqueous Suspension of Fullerene C-60 Nanoparticles in the Brains of Cyprinus carpio? *Environ. Sci. Technol.* **2009,** *43* (3), 948-953.
40. DeLoid, G. M.; Cohen, J. M.; Pyrgiotakis, G.; Demokritou, P., Preparation, characterization, and in vitro dosimetry of dispersed, engineered nanomaterials. *Nat. Prot.* **2017,** *12* (2), 355-371.
41. Taurozzi, J. S.; Hackley, V. A.; Wiesner, M. R., Ultrasonic dispersion of nanoparticles for environmental, health and safety assessment - issues and recommendations. *Nanotoxicology* **2011,** *5* (4), 711-729.
42. Taurozzi, J. S.; Hackley, V. A.; Wiesner, M. R., A standardised approach for the dispersion of titanium dioxide nanoparticles in biological media. *Nanotoxicology* **2013,** *7* (4), 389-401.
43. Drewes, C.; Ojea Jimenez, I.; Mehn, D.; Colpo, P.; Gioria, S.; Bogni, A.; Ponti, J.; Kinsner-Ovaskainen, A.; Gilliland, D.; Riego Sintes, J. *Physicochemical characterisation of gold, silica and silver nanoparticles in water and in serum-containing cell culture media, JRC TECHNICAL REPORT*; Publications Office of the European Union, 2017.
44. Ellison, S. L. R.; Barwick, V. J., Estimating measurement uncertainty: reconciliation using a cause and effect approach. *Accredit. Qual. Assur.* **1998,** *3* (3), 101-105.
45. Food and Drug Administration, Guidance for Industry: Bioanalytical Method Validation Revision 1. 2013; p 34.
46. Elliott, J. T.; Rosslein, M.; Song, N. W.; Toman, B.; Kinsner-Ovaskainen, A.; Maniratanachote, R.; Salit, M. L.; Petersen, E. J.; Sequeira, F.; Romsos, E. L.; Kim, S. J.; Lee, J.; von Moos, N. R.; Rossi, F.; Hirsch, C.; Krug, H. F.; Suchaoin, W.; Wick, P., Toward Achieving Harmonization in a Nanocytotoxicity Assay Measurement Through an Interlaboratory Comparison Study. *Altex-Alt. Anim. Exper.* **2017,** *34* (2), 201-218.
47. Hellack, B.; Nickel, C.; Albrecht, C.; Kuhlbusch, T. A. J.; Boland, S.; Baeza-Squiban, A.; Wohlleben, W.; Schins, R. P. F., Analytical methods to assess the oxidative potential of nanoparticles: a review. *Environ. Sci.: Nano* **2017,** *4* (10), 1920-1934.
48. Nel, A.; Xia, T.; Madler, L.; Li, N., Toxic potential of materials at the nanolevel. *Science* **2006,** *311* (5761), 622-627.
49. Arts, J. H. E.; Hadi, M.; Irfan, M. A.; Keene, A. M.; Kreiling, R.; Lyon, D.; Maier, M.; Michel, K.; Petry, T.; Sauer, U. G.; Warheit, D.; Wiench, K.; Wohlleben, W.; Landsiedel, R., A decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping). *Regul. Toxicol. Pharmacol.* **2015,** *71* (2), S1-S27.
50. Arts, J. H. E.; Irfan, M. A.; Keene, A. M.; Kreiling, R.; Lyon, D.; Maier, M.; Michel, K.; Neubauer, N.; Petry, T.; Sauer, U. G.; Warheit, D.; Wiench, K.; Wohileben, W.; Landsiedel, R., Case studies putting the decision-making framework for the grouping and testing of nanomaterials (DF4nanoGrouping) into practice. *Regul. Toxicol. Pharmacol.* **2016,** *76*, 234-261.
51. Ayres, J. G.; Borm, P.; Cassee, F. R.; Castranova, V.; Donaldson, K.; Ghio, A.; Harrison, R. M.; Hider, R.; Kelly, F.; Kooter, I. M.; Marano, F.; Maynard, R. L.; Mudway, I.; Nel, A.; Sioutas, C.; Smith, S.; Baeza-Squiban, A.; Cho, A.; Duggan, S.; Froines, J., Evaluating the toxicity of airborne particulate matter and nanoparticles by measuring oxidative stress potential - A workshop report and consensus statement. *Inhal. Toxicol.* **2008,** *20* (1), 75-99.
52. Rushton, E. K.; Jiang, J.; Leonard, S. S.; Eberly, S.; Castranova, V.; Biswas, P.; Elder, A.; Han, X. L.; Gelein, R.; Finkelstein, J.; Oberdorster, G., Concept of assessing nanoparticle hazards considering nanoparticle dosemetric and chemical/biological response metrics. *J. Toxicol. Environ. Health-Part A-Current Issues* **2010,** *73* (5-6), 445-461.

53. El Yamani, N.; Collins, A. R.; Runden-Pran, E.; Fjellsbo, L. M.; Shaposhnikov, S.; Zienolddiny, S.; Dusinska, M., In vitro genotoxicity testing of four reference metal nanomaterials, titanium dioxide, zinc oxide, cerium oxide and silver: towards reliable hazard assessment. *Mutagenesis* **2017,** *32* (1), 117-126.
54. Hanna, S. K.; Cooksey, G. A.; Dong, S.; Nelson, B. C.; Mao, L.; Elliott, J. T.; Petersen, E. J., Feasibility of using a standardized Caenorhabditis elegans toxicity test to assess nanomaterial toxicity. *Environ. Sci.: Nano* **2016,** *3* (5), 1080-1089.
55. Guadagnini, R.; Kenzaoui, B. H.; Walker, L.; Pojana, G.; Magdolenova, Z.; Bilanicova, D.; Saunders, M.; Juillerat-Jeanneret, L.; Marcomini, A.; Huk, A.; Dusinska, M.; Fjellsbo, L. M.; Marano, F.; Boland, S., Toxicity screenings of nanomaterials: challenges due to interference with assay processes and components of classic in vitro tests. *Nanotoxicology* **2015,** *9*, 13-24.
56. Horst, A. M.; Vukanti, R.; Priester, J. H.; Holden, P. A., An Assessment of Fluorescence- and Absorbance-Based Assays to Study Metal-Oxide Nanoparticle ROS Production and Effects on Bacterial Membranes. *Small* **2013,** *9* (9-10), 1753-1764.
57. Tarpey, M. M.; Wink, D. A.; Grisham, M. B., Methods for detection of reactive metabolites of oxygen and nitrogen: in vitro and in vivo considerations. *Amer. J. Physiol.-Regulat. Integrat. Compar. Physiol.* **2004,** *286* (3), R431-R444.
58. Cedervall, T.; Lynch, I.; Lindman, S.; Berggård, T.; Thulin, E.; Nilsson, H.; Dawson, K. A.; Linse, S., Understanding the nanoparticle–protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proceedings of the National Academy of Sciences* **2007,** *104* (7), 2050-2055.
59. Lundqvist, M.; Stigler, J.; Elia, G.; Lynch, I.; Cedervall, T.; Dawson, K. A., Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proc. Nat.Acad. Sci.***2008,** *105* (38), 14265-14270.
60. Bordeaux, J.; Welsh, A. W.; Agarwal, S.; Killiam, E.; Baquero, M. T.; Hanna, J. A.; Anagnostou, V. K.; Rimm, D. L., Antibody validation. *BioTechniques* **2010,** *48* (3), 197-209.
61. Piret, J. P.; Bondarenko, O. M.; Boyles, M. S. P.; Himly, M.; Ribeiro, A. R.; Benetti, F.; Smal, C.; Lima, B.; Potthoff, A.; Simion, M.; Dumortier, E.; Leite, P. E. C.; Balottin, L. B.; Granjeiro, J. M.; Ivask, A.; Kahru, A.; Radauer-Preiml, I.; Tischler, U.; Duschl, A.; Saout, C.; Anguissola, S.; Haase, A.; Jacobs, A.; Nelissen, I.; Misra, S. K.; Toussaint, O., Pan-European inter-laboratory studies on a panel of in vitro cytotoxicity and pro-inflammation assays for nanoparticles. *Arch. Toxicol.* **2017,** *91* (6), 2315-2330.
62. Glier, H.; Heijnen, I.; Hauwel, M.; Dirks, J.; Quarroz, S.; Lehmann, T.; Rovo, A.; Arn, K.; Matthes, T.; Hogan, C.; Keller, P.; Dudkiewicz, E.; Stüssi, G.; Fernandez, P., Standardization of 8-color flow cytometry across different flow cytometer instruments: A feasibility study in clinical laboratories in Switzerland. *J. Immunol. Meth.* **2017**.
63. Wu, D. Y.; Patti-Diaz, L.; Hill, C. G., Development and validation of flow cytometry methods for pharmacodynamic clinical biomarkers. *Bioanalysis* **2010,** *2* (9), 1617-1626.
64. Bohmer, N.; Rippl, A.; May, S.; Walter, A.; Heo, M. B.; Kwak, M.; Roesslein, M.; Song, N. W.; Wick, P.; Hirsch, C., Interference of Engineered Nanomaterials in Flow Cytometry: A Case Study. *Coll. Surf. B: Biointerfaces* **2018**.
65. Hande, K. R., Etoposide: Four decades of development of a topoisomerase II inhibitor. *Eur. J. Cancer* **1998,** *34* (10), 1514-1521.
66. Braafladt, S.; Reipa, V.; Atha, D. H., The Comet Assay: Automated Imaging Methods for Improved Analysis and Reproducibility. *Sci. Rep.* **2016,** *6*.
67. DaNa 2.0, Operating Instructions. (July 20, 2017), https://nanopartikel.info/en/nanoinfo/methods/992-operating-instructions
68. Karlsson, H. L.; Di Bucchianico, S.; Collins, A. R.; Dusinska, M., Can the Comet Assay be Used Reliably to Detect Nanoparticle-Induced Genotoxicity? *Environ. Mol. Mutag.* **2015,** *56* (2), 82-96.

69. Rajapakse, K.; Drobne, D.; Kastelec, D.; Marinsek-Logar, R., Experimental evidence of false-positive Comet test results due to TiO2 particle - assay interactions. *Nanotoxicology* **2013,** *7* (5), 1043-1051.
70. Lin, M. H.; Hsu, T. S.; Yang, P. M.; Tsai, M. Y.; Perng, T. P.; Lin, L. Y., Comparison of organic and inorganic germanium compounds in cellular radiosensitivity and preparation of germanium nanoparticles as a radiosensitizer. *Inter. J. Rad. Biol.* **2009,** *85* (3), 214-226.
71. Magdolenova, Z.; Lorenzo, Y.; Collins, A.; Dusinska, M., Can standard genotoxicity tests be applied to nanoparticles? *J. Toxicol. Environ. Health-Part A-Current Issues* **2012,** *75* (13-15), 800-806.
72. Ferraro, D.; Anselmi-Tamburini, U.; Tredici, I. G.; Ricci, V.; Sommi, P., Overestimation of nanoparticles-induced DNA damage determined by the comet assay. *Nanotoxicology* **2016,** *10* (7), 861-870.
73. George, J. M.; Magogotya, M.; Vetten, M. A.; Buys, A. V.; Gulumian, M., An Investigation of the Genotoxicity and Interference of Gold Nanoparticles in Commonly Used In Vitro Mutagenicity and Genotoxicity Assays. *Toxicol. Sci.* **2017,** *156* (1), 149-166.
74. Gerloff, K.; Albrecht, C.; Boots, A. W.; Forster, I.; Schins, R. P. F., Cytotoxicity and oxidative DNA damage by nanoparticles in human intestinal Caco-2 cells. *Nanotoxicology* **2009,** *3* (4), 355-364.
75. Organization for Economic Cooperation and Development, ENV/JM/MONO(2015)17/PART4/ANN2: Dossier on Titanium Dioxide - Part 4 - NM 102 Annex 2. Paris, France, 2015.
76. Ellison, S. L. R., Uncertainties in qualitative testing and analysis. *Accredit. Qual. Assur.* **2000,** *5*, 346-348.
77. Ellison, S. L. R.; Gregory, S.; Hardcastle, W. A., Quantifying uncertainty in qualitative analysis. *Analyst* **1998,** *123* (5), 1155-1161.
78. Drasler, B.; Kucki, M.; Delhaes, F.; Buerki-Thurnherr, T.; Vanhecke, D.; Korejwo, D.; Chortarea, S.; Barosova, H.; Hirsch, C.; Petri-Fink, A.; Rothen-Rutishauser, B.; Wick, P., Single exposure to aerosolized graphene oxide and graphene nanoplatelets did not initiate an acute biological response in a 3D human lung model. *Carbon* **2018,** *137*, 125-135.

Figures

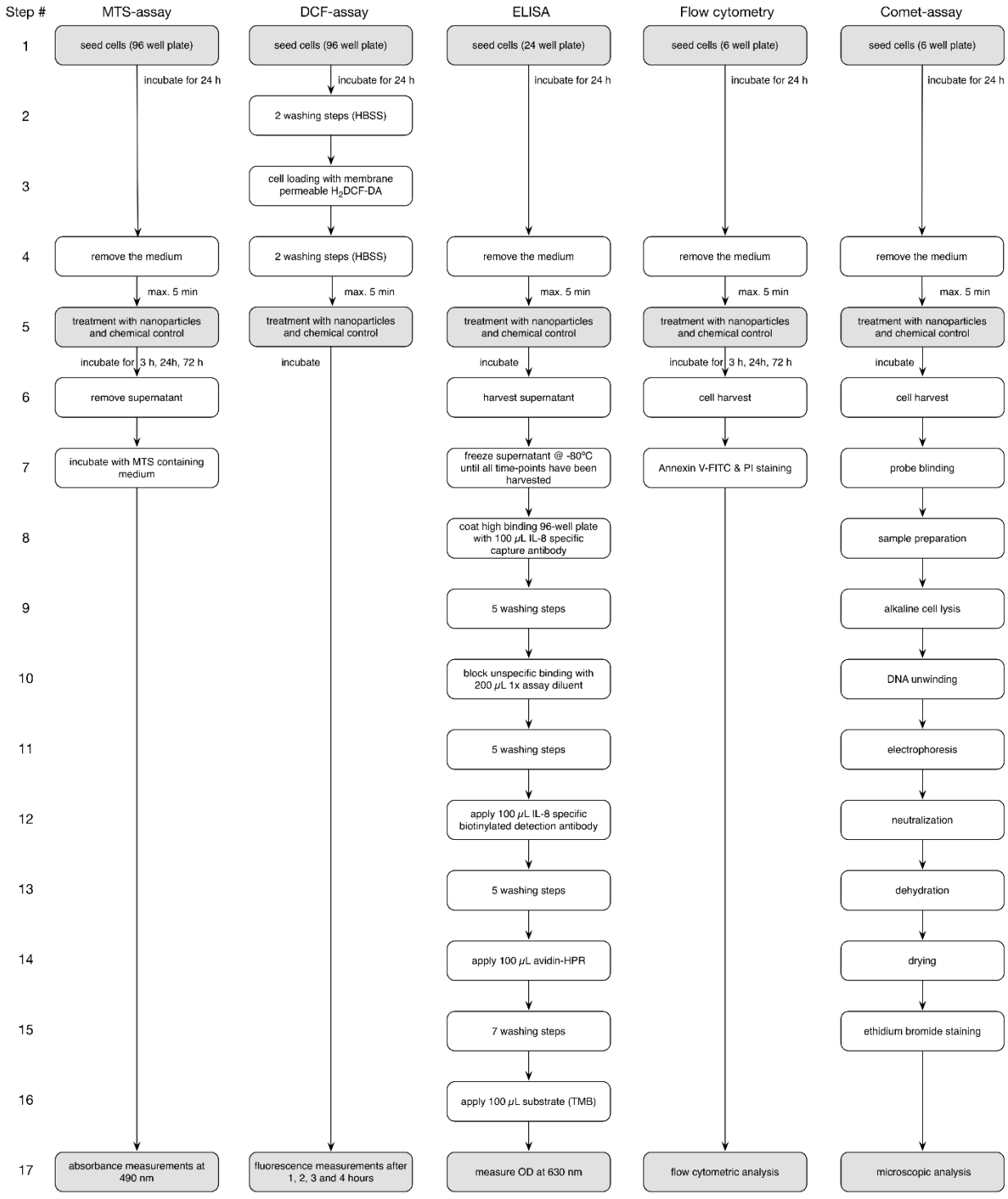
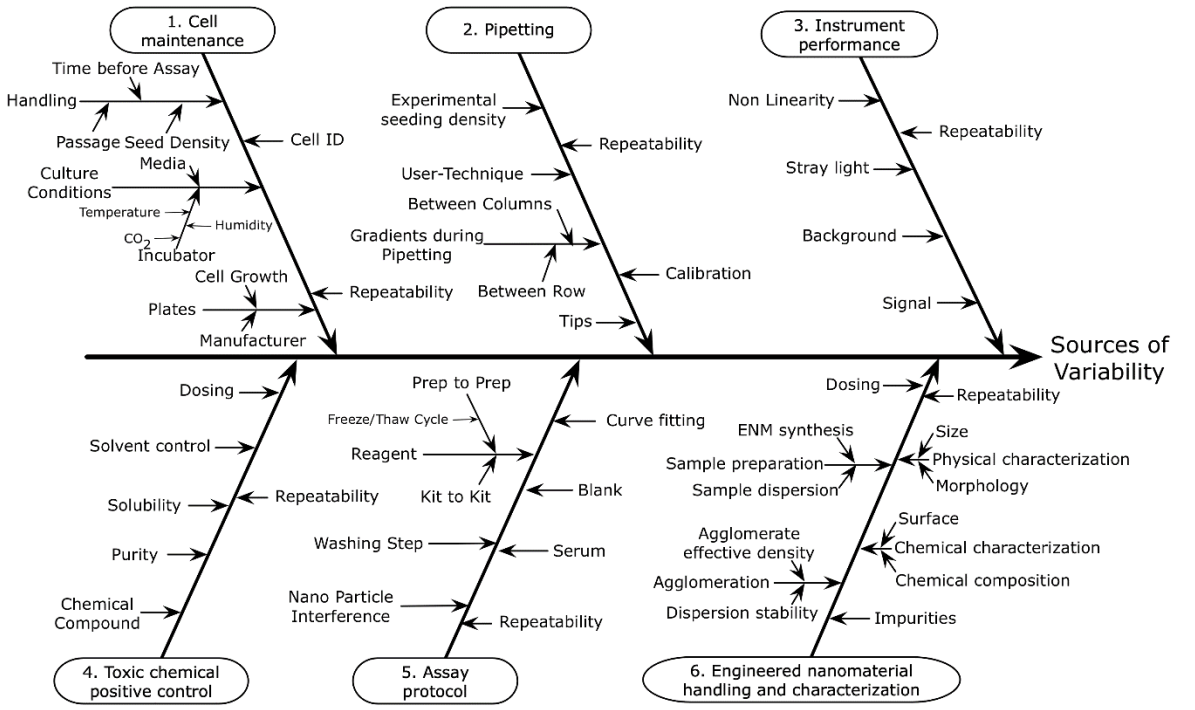
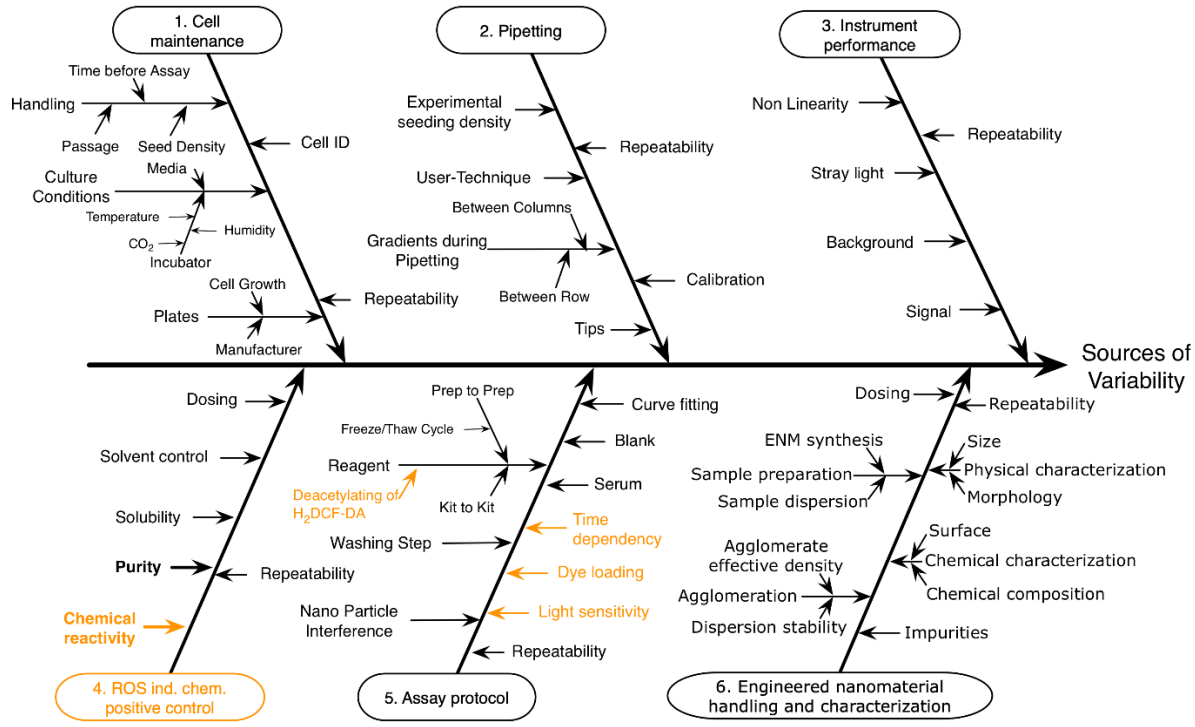


Figure 1: Flow charts illustrating steps for MTS, DCF, ELISA, flow cytometry, and Comet assays.



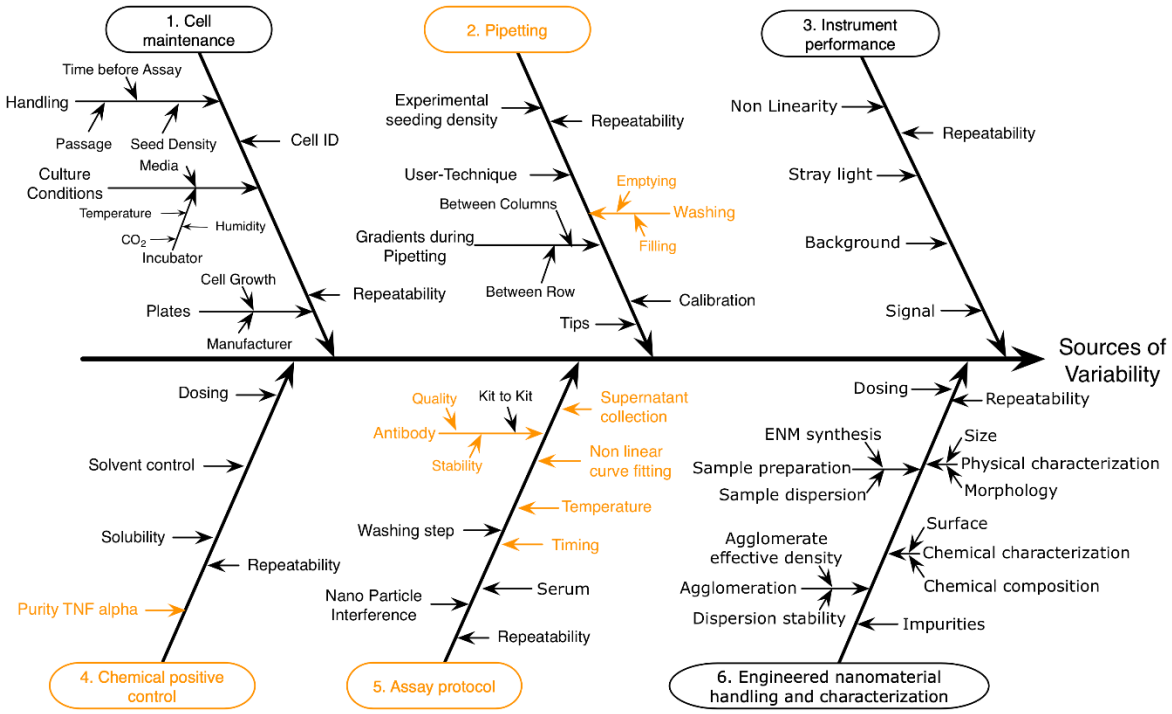
791

792 Figure 2: Cause-and-effect diagram of MTS assay. Modified and reprinted with permission from the
793 American Chemical Society.[15]



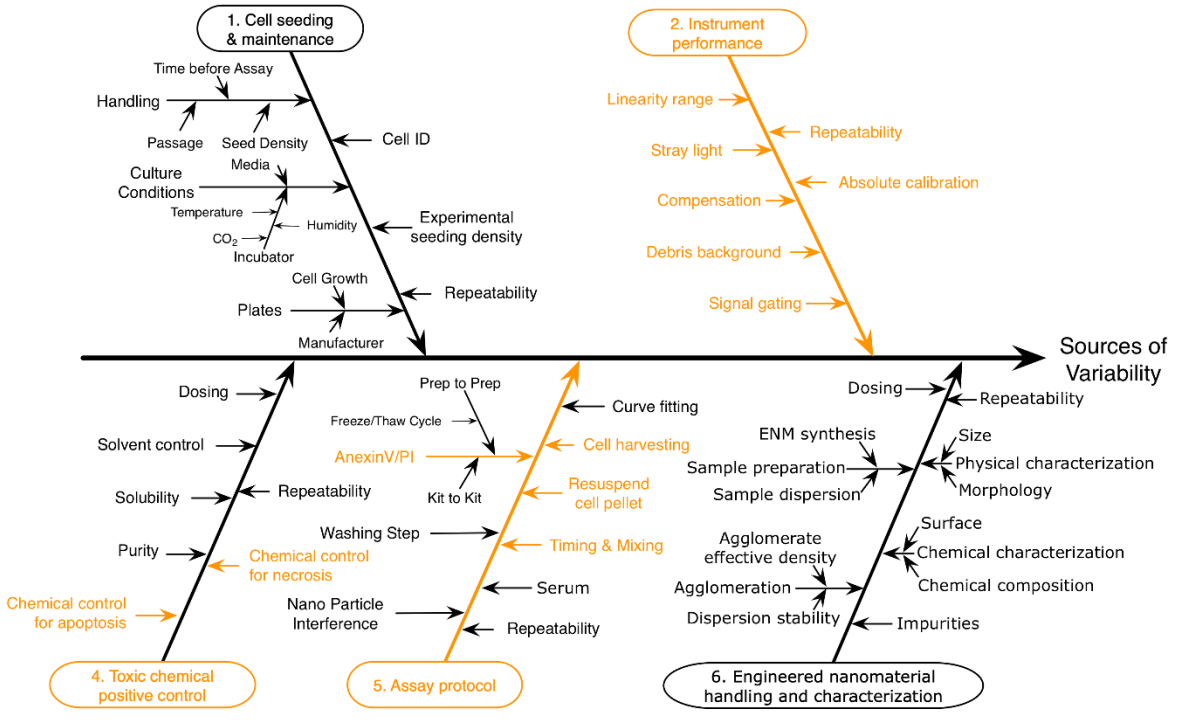
794

795 Figure 3: Cause-and-effect diagram of the DCF assay. Parts of the diagram in orange font indicate differ
796 from the cause-and-effect diagram for the MTS assay.



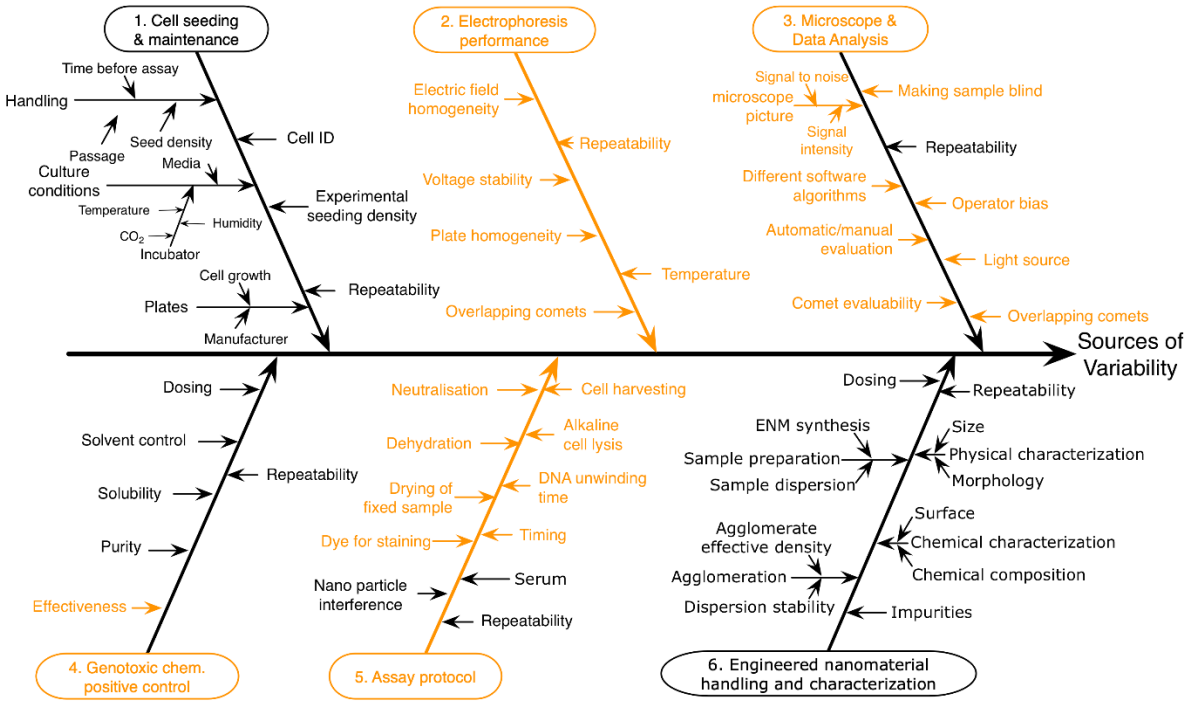
797

798 Figure 4: Cause-and-effect diagram of the ELISA for analyzing IL-8. Parts of the diagram in orange font
799 indicate differ from the cause-and-effect diagram for the MTS assay.



800

801 Figure 5: Cause-and-effect diagram for the flow cytometry for measuring necrotic and apoptotic cells.
802 Parts of the diagram in orange font indicate differ from the cause-and-effect diagram for the MTS assay.



803

804 Figure 6: Cause-and-effect for analyzing DNA damage of cells exposed to ENMs with the Comet assay.
805 Parts of the diagram in orange font indicate differ from the cause-and-effect diagram for the MTS assay.

806

807 Table 1: Overview for each of the five assays

Assay	Assay purpose	What is measured
MTS assay	Measures metabolical activity as an indirect estimate of the amount of living cells	The amount of metabolically active cells (i.e., living/healthy cells) is estimated by spectrophotometrically quantifying the amount of MTS reagent metabolized
DCF	Measures reactive oxygen species in cellular systems	A cell-permeable, non-fluorescent dye is added to cells, the cells cleave off the diacetate moiety, the dye is reduced by ROS, and the reduced dye is detected using fluorescence
ELISA Method with IL-8	Measures the cellular release of the cytokine Interleukin-8 (IL-8)	After cellular exposure to a potential stimulus (e.g., ENM), the supernatant is removed and added to a plate which has an antibody attached surface; a second biotinylated antibody is then added to bind to the same cytokine but at a different epitope, and lastly the cytokine concentration is measured spectrophotometrically after horseradish peroxidase linked to avidin reacts using tetramethylbenzidine as the substrate reacts with the biotinylated antibody
Annexin V/ Propidium Iodide Flow Cytometry	Measures cell viability and necrotic/apoptotic cells	Dyes are added to distinguish between viable, necrotic, and apoptotic cells and the cells are measured using flow cytometry
Comet assay	Measures DNA damage	After cell lysis and single cell gel electrophoresis the addition of a DNA-intercalating dye and microscopic imaging allow quantifying DNA breaks in each cell; the relative fluorescence intensity in the tail (% DNA in tail), the tail length as well as the tail moment (taking tail intensity as well as tail length into consideration) are accepted metrics for DNA damage quantification

808

809 Table 2: Control preliminary experiments to assess sources of variability and bias for the DCF assay with
810 ENMs

Source of variability/bias	Step #	Potential control measurements
Positive chemical control	5	Evaluate potential positive chemical controls ideally to find a compound that always induces the same amount of ROS
Time dependency	17	Perform kinetic measurement because only a small fraction of ENMs may have reached the cells at the bottom of the test plate at early time points (e.g., 4 h). However, it is also valuable to obtain data about the amount of ROS produced shortly after ENM exposure with cells to try to measure the initial release of ROS. It is also possible to quantify ENM uptake in separate experiments prior to the DCF assay (although in practice this may be challenging to measure for some ENMs).
False positive results from extracellular production of ROS	5	Cleave the DA from H_2DCF-DA in an acellular experiment and then investigate for potential production of DCF signal during incubation with ENMs as a result of extracellular ROS production. If this is observed, there could be a false positive signal in cellular assays if both extracellular deacetylation and ROS production occur. In addition, if such a signal is observed it could further indicate, that the ENM's reactive surface processes the H_2DCF dye without production of ROS.

811

812

Table 3: Control preliminary experiments to assess sources of variability and bias for the ELISA protocol to quantify IL-8 with ENMs

Source of variability/bias	Step #	Potential control measurements
Large number of pipetting/washing steps may increase variability	9, 11, 12, 15	Perform experiments where the washing steps are performed manually or if this step is automated to evaluate the improvement in performance of the assay (e.g., decreased variability of negative control cells and cells exposed to the positive chemical control).
ENM interference with the assay through inherent ENM absorbance or adsorption of assay components	5	Spike ENM into different steps of the assay procedure in the presence and absence of the antigen (used for standard curve preparation) according to SOP on DaNa webpage (V.I.G.O. (2014) "NM interference in an enzyme-linked immunosorbent assay (ELISA) v1.0") [67] and evaluate the impact.[78]
The kinetics of the color reaction may require optimization	17	Take data at several points after substrate application (5, 10, 15, or 20 min) to determine the optimal protocol parameters
Instrument Performance	17	Absorbance standards to characterize function of the instrument; measure with an alternative instrument for comparison
Adsorption of cytokines onto ENMs decreases response	5	A control experiment could be performed to assess adsorption of cytokines by the test ENM. This would consist of incubating the ENMs with a known concentration of IL-8 for a certain interval and then conducting the ELISA method to assess the IL-8 recovery.
Batch to batch variability of ELISA reagents	Not applicable	Run ELISA on the same plate with old and new reagents (standard curve only)

818 Table 4: Control preliminary experiments to assess sources of variability and bias for the Annexin PI flow
819 cytometry assay with ENMs

Source of variability/bias	Step #	Potential control measurements
Identify a concentration and exposure duration of the positive chemical control to produce a reproducible amount of apoptosis and necrosis cells	5	Conduct preliminary experiments to assess several concentrations and time points. It would be helpful also to test a reference material with a reproducible number of apoptotic and necrotic cells for benchmarking assay performance and comparing to test results.
Interference of ENMs with instrument detection	17	Conduct the following pre-experiments; 1) flow cytometric analysis of cell-free samples (i.e. ENM only) to assess ENMs presence and their interaction with staining solutions; 2) flow cytometric analysis of untreated control cells with ENMs spiked in to assess if ENMs alter the detection of unstained/stained cells or interact with cell debris even if the ENMs are not detected themselves; 3) use other approach to detect fluorescence spectrum of ENMs and ENMs with staining reagents (e.g., test with a plate reader or during physicochemical characterization of the ENMs)

820

821

822 Table 5: Control preliminary experiments to assess sources of variability and bias for the Comet assay
823 with ENMs

Source of variability/bias	Step #	Potential control measurements
ENM interference with DNA migration	5	Add ENM after cell lysis and assess if there is a change in the Comet size; analyze the Comet tail to assess if ENMs are present such as using electron microscopy or hyperspectral imaging. Assess if washing or separation procedures can be used to minimize ENM concentration in the Comet tail.
Variable voltage during electrophoresis within the plate could influence DNA migration	11	Rotate the slides of control samples to assess if there are differences.
Artifactual light damage after conclusion of cell experiment from photoactive ENMs	5	Conduct the cell exposure as usual and then conduct the remaining steps of the assay with some samples processed with laboratory light and others processed in the dark. Assess if there is a difference between samples processed in the light or dark.

824

825

Author biographies

Elijah J. Petersen: Elijah completed his PhD at the University of Michigan in Environmental Engineering. Then, he completed postdocs at the University of Joensuu (Finland) on a Fulbright scholarship and then the University of Michigan before joining NIST as a National Research Council postdoctoral fellow. He became a staff scientist at NIST in 2010 and works in the Cell Systems Science group in the Biosystems and Biomaterials division. His research currently focuses on the development of robust, reproducible *in vitro* test methods. He is an associate editor for Nanotoxicology and Nanoimpact and on the editorial board of Environmental Pollution and Environmental Toxicology and Chemistry.

Cordula Hirsch: Cordula Hirsch is a scientist and project leader at Empa developing *in vitro* methods according to the 3Rs principles (replacement, reduction and refinement of animal experiments). A special focus is the improvement and adaptation of *in vitro* cytotoxicity methods for nanosafety as well as nanomedical purposes and to achieve reliable, reproducible and robust test results. She studied biology at the Universities of Konstanz and Freiburg (Germany) and received her PhD in 2007 on cell signaling in the developing and adult nervous system of mice. In 2008, she joined the Particles-Biology interactions laboratory at Empa as a Postdoc.

John T. Elliott: John graduated in 1990 with a bachelor's degree in Physics from the University of Massachusetts. After serving as a technical assistant at the Massachusetts General Hospital, he then received a PhD in Physiology and Biophysics from SUNY at Stony Brook in 1999 after which he joined the National Institute of Standards and Technology as a National Research Council postdoctoral fellow. He became a staff scientist at NIST in 2001 and now is the group leader for the Cell Systems Science group in the Material Measurement Laboratory's Biosystems and Biomaterials division.

Harald F. Krug: Harald Krug studied Chemistry and Biology. He received his Ph.D. from the Georg-August University in Göttingen for his work in animal physiology. After a postdoc period at the Helmholtz Centre Munich he took over the Department for Environmental Toxicology at the Karlsruhe Institute of Technology. Since 2007, he teaches at the University Berne. At the Swiss Federal Institute for Materials Science & Technology (Empa) he headed the laboratory for Nanomaterials-Biology Interaction and was member of the board from 2010 to 2014. Since 2017 he is retired and is busy with his own company NanoCASE, which he founded in 2014.

Leonie Aengenheister: Dr. Leonie Aengenheister received her PhD from ETH Zurich, Switzerland, in 2018 for establishing a novel advanced *in vitro* placental model and revealing new insights on placental uptake and translocation of different nanomaterials. She then worked as a postdoctoral researcher at the University of Texas Medical Branch in Galveston, where she designed nanoparticle-based drug carriers destined for therapies during pregnancy. Currently, she continues her investigations on nanomaterial-placenta interactions and potential associated effects on maternal and fetal health as a postdoctoral researcher at Empa - the Swiss Federal Laboratories for Materials Science and Technology in St. Gallen, Switzerland.

Ali Talib Arif: Dr. Arif is working as a postdoctoral researcher in the group Environmental Toxicology and Nanotoxicology, Institute for Infection Prevention and Hospital Epidemiology, University Medical Center Freiburg, Germany. He is also a guest scientist at the Kurdistan Institution for Strategic Studies and Scientific Research, Iraq. He holds both B.S and M.Sc. degrees in Biology from the University of Hohenheim, Germany. Dr. Arif's current research studies are deal with molecular mechanisms of environmental chemicals and fine or ultrafine particles toxicity that play a major role to understanding the Health effects associated with particles exposure, so he has published three papers on these topics.
Alessia Bogni: Alessia Bogni received a degree in Molecular Biology in 1998 at the University of Milan (Italy), and a PhD in toxicology and pharmacology from University of Konstanz (Germany) in 2003. She gained experience in the field working in Pharmacia&Upjoin (Milan, Italy) and at St. Jude Children Research Hospital (Memphis, TN, USA). She started working at the European Commission Joint Research Centre (Ispra-Italy) in 2006.

Agnieszka Kinsner-Ovaskainen: Agnieszka Kinsner-Ovaskainen obtained her M.D. degree from the Medical University of Warsaw, Poland and the Ph.D. degree in Life Sciences from the University of Konstanz, Germany. She joined the European Commission Joint Research Centre (JRC) as official in 2007. Currently she is working in the Health in Society Unit (JRC), responsible for the Central Registries and the European-level coordination of two networks of population-based registries for the surveillance of congenital anomalies (EUROCAT) and cerebral palsy (SCPE).

Sarah May: Sarah May studied biological sciences at the University of Konstanz (Germany) and obtained her bachelors degree in 2011, followed by her masters degree in 2014. During her PhD she investigated the effect of different engineered nanomaterials on DNA damage and repair pathways at Empa, St. Gallen in collaboration with the University of Konstanz. In 2018 she received her PhD from the University of Konstanz (Germany). Currently, she is working as medical affairs manager in the medical device industry in Winterthur, Switzerland.

Tobias Walser: Dr. Tobias Walser received his PhD in Environmental Engineering on Life Cycle based Risk Assessment of Nanomaterials. After a postdoctoral stay at the U.S. Environmental Protection Agency on modelling the toxic impact of nanomaterials, he worked for the Swiss Federal Office of Public Health. Responsibilities were regulatory risk assessment of nanomaterials, and expert work in various international gremia. In 2017, Tobias Walser founded the company "Vereala", which offers expert knowledge on emerging technologies with a focus on nanomaterials

Peter Wick: Peter Wick heads since 2014 the research laboratory for Particles-Biology Interactions at the Federal Laboratories on Materials Science and Technologies Empa in St. Gallen. His general research interest is to study the interactions of nanomaterials with human tissues *in vitro* and *ex vivo* with the purpose to obtain detailed mechanistic understanding about their uptake, accumulation, transport and effect on different types of cells or entire tissue. He is a member of the advisory board of the Swiss Action Plan on Nanomaterials, member of the EDQM working group for NBCs, Editorial Board Member of Nanotoxicology and associated editor of Journal NanoImpact.

Matthias Roesslein: Matthias Rösslein studied chemistry at the university of Basel, where he also received his PhD degree in 1989. Afterward he spent to 2 years as a postdoc at the university of Chicago

and 4 years as an assistant professor at the physical-chemical institute of the university Zurich, before joining Empa in 1996 as a leading expert in the 'evaluation of measurement uncertainty and metrology'. In 2006 he was appointed a position as 'Senior Scientist' and changed to the "Particles-Biology Interaction" laboratory focusing on the standardization of *in vitro* assays to elucidate the effect of engineered nanomaterials on different cell types.