Toxicology Mechanisms and Methods

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/itxm20

Development of a standard operating procedure for the $DCFH_2$-DA acellular assessment of reactive oxygen species produced by nanomaterials

Matthew Boyles, Fiona Murphy, William Mueller, Wendel Wohlleben, Nicklas Raun Jacobsen, Hedwig Braakhuis, Anna Giusti & Vicki Stone

To cite this article: Matthew Boyles, Fiona Murphy, William Mueller, Wendel Wohlleben, Nicklas Raun Jacobsen, Hedwig Braakhuis, Anna Giusti & Vicki Stone (2022): Development of a standard operating procedure for the $DCFH_2$-DA acellular assessment of reactive oxygen species produced by nanomaterials, Toxicology Mechanisms and Methods, DOI: 10.1080/15376516.2022.2029656

To link to this article: https://doi.org/10.1080/15376516.2022.2029656

TOXICOLOGY MECHANISMS AND METHODS
https://doi.org/10.1080/15376516.2022.2029656

Taylor & Francis
Taylor & Francis Group

RESEARCH ARTICLE

OPEN ACCESS Check for updates

Development of a standard operating procedure for the $DCFH_2$-DA acellular assessment of reactive oxygen species produced by nanomaterials

Matthew Boyles[a]*, Fiona Murphy[b]*, William Mueller[a], Wendel Wohlleben[c], Nicklas Raun Jacobsen[d], Hedwig Braakhuis[e], Anna Giusti[f] and Vicki Stone[b]

[a]Institute of Occupational Medicine (IOM), Edinburgh, UK; [b]NanoSafety Group, Heriot-Watt University, Edinburgh, UK; [c]Department of Material Physics and Department of Experimental Toxicology & Ecology, BASF SE, Ludwigshafen, Germany; [d]National Research Centre for the Working Environment (NFA), Copenhagen, Denmark; [e]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands; [f]Department of Chemical and Product Safety, German Federal Institute for Risk Assessment (BfR), Berlin, Germany

ABSTRACT

Improved strategies are required for testing nanomaterials (NMs) to make hazard and risk assessment more efficient and sustainable. Including reduced reliance on animal models, without decreasing the level of human health protection. Acellular detection of reactive oxygen species (ROS) may be useful as a screening assay to prioritize NMs of high concern. To improve reliability and reproducibility, and minimize uncertainty, a standard operating procedure (SOP) has been developed for the detection of ROS using the 2′,7′-dichlorodihydrofluorescein diacetate ($DCFH_2$-DA) assay. The SOP has undergone an inter- and intra-laboratory comparison, to evaluate robustness, reliability, and reproducibility, using representative materials (ZnO, CuO, Mn_2O_3, and $BaSO_4$ NMs), and a number of calibration tools to normalize data. The SOP includes an NM positive control (nanoparticle carbon black (NPCB)), a chemical positive control (SIN-1), and a standard curve of fluorescein fluorescence. The interlaboratory comparison demonstrated that arbitrary fluorescence units show high levels of partner variability; however, data normalization improved variability. With statistical analysis, it was shown that the SIN-1 positive control provided an extremely high level of reliability and reproducibility as a positive control and as a normalization tool. The NPCB positive control can be used with a relatively high level of reproducibility, and in terms of the representative materials, the reproducibility CuO induced-effects was better than for Mn_2O_3. Using this $DCFH_2$-DA acellular assay SOP resulted in a robust intra-laboratory reproduction of ROS measurements from all NMs tested, while effective reproduction across different laboratories was also demonstrated; the effectiveness of attaining reproducibility within the interlaboratory assessment was particle-type-specific.

ARTICLE HISTORY
Received 27 August 2021
Revised 29 October 2021
Accepted 8 January 2022

KEYWORDS
Nanomaterials; nanoparticles; nanoforms; reactive oxygen species; oxidative stress; free radicals

Introduction

Nanotechnology is a key enabling technology of the 21st century, with the exploitation of nanomaterials (NMs) in a range of applications associated with many benefits for society and the global economy. However, uncertainties regarding the possible risks posed by the increasing number of NMs and their diverse nanoforms (NFs) to human health and the environment need to be better understood in order to facilitate the success of this technology. With the continued advancement of the 3R's principle, advocating Replacement, Reduction, and Refinement of animal studies, suitable and reliable alternatives to animal testing are required for the sustainable safety assessment of NMs in the future. In fact, there is a legal obligation under REACH, EC No 1907/2006 (EC 2006) for companies to pursue alternative testing strategies in lieu of data obtained from the use of animal models. However, the use of 'New Approach Methodologies (NAMs)' in safety assessment requires robust and reliable *in vitro* and *in chemico* tools to enable confidence in the toxicological data obtained (ECHA 2016).

Intrinsic reactive oxygen species (ROS) production by NMs represents a key mechanism by which certain NMs cause cellular toxicity via oxidative stress. Measurement of intrinsic ROS production in a cell-free environment may be used as a screening tool to identify NMs which may be considered to be biologically reactive and require further toxicity testing. However, standardization of common techniques used to detect ROS production by NMs is lacking and has led to the generation of conflicting results and uncertainty in the utility of these approaches for assessing potential hazards.

The lipophilic non-fluorescent molecule 2′-7′-dichlorodihydrofluorescein diacetate ($DCFH_2$-DA) is a commonly used tool to study cellular and acellular ROS produced in response

CONTACT Matthew Boyles matthew.boyles@iom-world.org Institute of Occupational Medicine (IOM), Research Avenue North, Riccarton, Edinburgh, EH14 4AP, UK
*Contributed equally to this work.

Supplemental data for this article can be accessed here.

to substances including combustion-derived particles (Wilson et al. 2002; Wilson et al. 2007) and engineered NMs (Foucaud et al. 2007; Rothen-Rutishauser et al. 2010; Sauvain et al. 2013), and as such it is paramount that a robust Standard Operating Procedure (SOP) is made available for this method. The assay is based on measuring the oxidation of the non-fluorescent probe 2′,7′-dichlorodihydrofluorescein ($DCFH_2$), which when exposed to ROS yields the highly fluorescent 2′,7′-dichlorofluorescein (DCF). $DCFH_2$-DA was developed for use in oxidation experiments as a replacement for fluorescin, which is relatively instable and can be oxidized to its fluorescent form fluorescein. The improved stability in long term storage obtained by acetylation (Brandt and Keston 1965) did not influence the fluorescence properties of the probe, and being an acetyl group could easily be removed by hydrolysis in response to an alkaline solution or enzyme activity (Brandt and Keston 1965). Later, this probe was adopted in the assessment of ROS generated intracellularly (Bass et al. 1983), with the diacetate moiety allowing entry into cells, and subsequent cleavage of the diacetate by cytosolic esterases (Bass et al. 1983).

Due to the high sensitivity of the probe to oxidants, the $DCFH_2$-DA assay is not specific to any one ROS and allows for detection of a vast range of ROS/RNS (reactive nitrogen species), such as $RO^{\cdot}{}_2$, $RO^{\cdot}$, $OH^{\cdot}$, HOCl, and ONOO (Doak et al. 2009). For some free radicals, however, there is an incongruence in opinions, singlet oxygen ($^1O_2$), for example, has been reported to oxidize $DCFH_2$ (Daghastanli et al. 2008) and conversely, to not react with $DCFH_2$ at all (Bilski et al. 2002). The non-discriminatory nature of the $DCFH_2$ probe has led to its frequent use *in vitro* and in cell-free settings, determining ROS production by NMs such as carbon nanotubes, fine and ultra-fine carbon black, titanium dioxide, zinc oxide, silver, silver oxide, gold, copper phthalocyanine and copper oxide (Karlsson et al. 2008; Rothen-Rutishauser et al. 2010; Pal et al. 2012; Aranda et al. 2013; Schlinkert et al. 2015; Hellack et al. 2017; Pang et al. 2017). In a comparison of three acellular tests for assessing the oxidation potential of NMs, Sauvain et al. (2013) found that ascorbic acid assay (AA-assay) was not sufficiently sensitive nor robust enough to detect ROS from carbonaceous and metal/metal oxide (Me/MeOx) NMs. In addition, they found the $DCFH_2$ assay was approximately 10 times more sensitive than the dithiothreitol (DTT) assay in detection of intrinsic ROS generation. Moreover, the $DCFH_2$ assay is quick, inexpensive and does not require specialized equipment or technical expertise, which should facilitate the potential wide-spread adoption of this method.

Consensus on the appropriateness of the assay for detection of ROS by NM has been hindered by uncertainty in results due to both inherent limitations of the assay itself and differences in protocols leading to conflicting results. One major problem inherent to $DCFH_2$ assay use is its unstable nature. The $DCFH_2$ dye is slowly oxidized to the fluorescent DCF species in the air and is also prone to photo-oxidation by the laser light utilized for fluorescence excitation. Thus, DCF detection can generate false-positive results (Sarvazyan 1996; Wang and Joseph 1999) and background values increase with time. Furthermore, there is a lack of harmonization between protocols commonly used. Variations amongst procedures relate to the use of oxidation enhancing materials (e.g. horseradish peroxidase) to mimic cellular conditions, and differences in incubation time, the concentration of probe, particle preparation, suspension buffer used, sonication conditions, and fluorescent wavelength (Foucaud et al. 2007; Pal et al. 2012). Coupled with the $DCFH_2$ sensitivity to auto-oxidation, differences in these variables across protocols will heavily influence their results, making interpretation and comparisons with other literature more difficult. Lack of quantitative representation of the $DCFH_2$-DA assay results also hinders the potential to compare data between laboratories as results are most commonly represented as arbitrary fluorescent units (AU). This issue could be potentially resolved with the use of reliable methods for data normalization, such as the adoption of a chemically stable fluorescein dose-response curve, or the determination of reliable positive controls. Either approach could facilitate the quantitative reporting of results and help correlate the DCF signal generated in different laboratories, or to compare studies conducted over time.

The $DCFH_2$-DA assay, like most of the other conventional colorimetric (dye-based) assays that have been used for NM toxicity evaluations, was originally standardized and optimized for chemical compounds and has often been adopted with little modifications and method evaluation of the potential for NM interference. For example Doak et al. (2009) reported that dextran-coated iron oxide NMs were able to interfere with the fluorescence emission of DCF, the degree of which varied according to the concentration of the dye and the oxidation state of the NMs. The potential for interference by the test NMs will need to be mitigated to increase confidence in the assay results.

The limitations of the $DCFH_2$-DA assay have previously been addressed and basic assay conditions, such as selection of suspension media, or operational parameters such as sonication method have been identified as highly influential to the assay results (Pal et al. 2012), highlighting the need to standardize the assay. Building on this previous work, this paper describes an SOP that has been developed for a commonly used method for the detection of ROS via the $DCFH_2$-DA assay. The SOP has undergone an inter- and intra-laboratory comparison, which has included repeat testing on either side of optimization procedures. We transparently describe the testing of assay parameters which led to the development of a provisional SOP, which underwent further refinement during interlaboratory comparison of ROS generation of a panel of representative NMs. Numerous procedures, pertaining to data handling processes and assay controls, have been employed to systemically assess which methods are optimum for obtaining good reproducibility across different laboratories, and with different operators.

Materials and methods

Development of standard operating procedure

Parameters tested before finalizing the SOP

During the development of the SOP a number of parameters were evaluated to ensure the optimal conditions were used and critical factors influencing the $DCFH_2$-DA assay were addressed within the SOP. The details of these are recorded within supplemental material and included: selection of appropriate positive controls, suitability of $DCFH_2$-DA concentrations, the impact of proteins present in exposure medium, the effect of storage of fluorescein diacetate (F-DA) on fluorescence signal, the effect of batch preparation on fluorescein fluorescence reproducibility, assessment of fluorescence quenching and enhancement due to NMs, and the impact of plate location bias or cross-well contamination on fluorescence signals.

Final SOP

The full SOP is provided as supplemental material, and given only briefly here.

Chemicals and materials. 3-Morpholinosydnonimine hydrochloride (SIN-1 hydrochloride) was sourced from Abcam; the representative materials used included NMs of zinc oxide (ZnO, JRCNM01100a, formally NM-110) and barium sulfate ($BaSO_4$, NM-220) which were obtained from the European Joint Research Center (JRC, Ispra, Italy) (JRC), manganese oxide (Mn_2O_3) from Skyspring Nanomaterials, and copper oxide nanopowder (ca. 15 nm) (CuO) from PlasmaChem (Berlin, Germany); Printex 90 nanoparticle carbon black (NPCB) from Evonik/Degussa was used as a NM positive control; all other reagents/chemicals were sourced from Sigma.

$DCFH_2$-DA assay. Hydrolysis of $DCFH_2$-DA to $DCFH_2$ was performed with incubation of 200 μM $DCFH_2$-DA in 0.01 M NaOH for 30 min at room temperature, protected from light, followed by the addition of 0.1 M PBS solution (pH7.4) to stop the deacetylation. This provided 50 μM $DCFH_2$, which was further diluted to 10 μM with 0.01 M PBS. This solution was kept on ice until use. The fluorescein diacetate (F-DA) standard curve was prepared with incubation of 200 μM F-DA in 0.01 M NaOH for 5 min at room temperature, protected from light. The reaction was stopped using 0.1 M PBS, generating a fluorescein concentration of 50 μM, this was further diluted with 0.01 M PBS to 1 μM followed by serial dilutions to achieve the standard curve of 0.001, 0.004, 0.012, 0.037, 0.111, 0.333, 1.000 μM.

The representative materials, positive NM control (NPCB) and positive chemical control (SIN-1) were prepared in phenol red-free medium (MEM), either in the absence or presence of 2% fetal calf serum (FCS). Test particles and NPCB were suspended at 1 mg/ml, vortexed, and then ultra-sonicated in a sonicating water bath for 15 min prior to dilution to the required highest concentration (125 μg/ml for NPCB, while test particles remained at 1 mg/ml); SIN-1 (stored in aliquots of 200 mM in H_2O) was diluted to 100 μM in MEM. Samples were vortexed before use, and further dilutions were made using a 96-well, round-bottomed 'loading' plate as depicted in the supplemental material, providing the following concentrations for test materials: 16, 31, 63, 125, 250, 500, 1000 μg/ml, for NPCB: 16, 31, 63, 125 μg/ml, and for SIN-1: 12.5, 25, 50, 100 μM. These provided final assay concentrations of 1.6, 3.1, 6.3, 12.5, 25, 50, 100 μg/ml for test materials, 1.6, 3.1, 6.3, 12.5 μg/ml for NPCB, and 1.3, 2.5, 5, 10 μM for SIN-1.

All reagents were added to a black clear-flat-bottom 96-well plate (plate layout is provided in the supplemental material) in triplicate. A volume of 250 μl of each fluorescein standard curve concentration was added first, followed by 25 μl of the no treatment (medium only), positive controls, and test materials using a multi-channel pipette from the loading plate, followed by 225 μl $DCFH_2$ reaction mix. The samples in the plate were then read immediately at ex/em 485/530, and then after 30, 60, and 90 min; between measurements, the plate was stored at 37 °C, protected from light. The assay was conducted with at least three independent replicates.

Particle interference. Any dose-dependent auto-fluorescence or fluorescence quenching (or enhancement) was assessed using PBS or an F-DA solution, respectively. The particles were prepared as described above and added to a black clear-bottom 96-well plate (plate layout is provided in the supplemental material) in triplicate, instead of adding $DCFH_2$, 225 μl of either PBS or 0.1 μM F-DA was added; the sample in the plate were then read immediately using ex/em 485/530.

Refinement of SOP between first and second interlaboratory comparison. The method used in the two independent interlaboratory comparison studies was similar, however in the second study a strict order of reagent preparation was used, and a common strategy for particle dilution and handling was devised; the SOP provided in the supplemental material is the final SOP, as used in the second interlaboratory comparison study.

Data analysis

A data analysis template is provided as supplemental material, which includes considerations of the test for particle interference in the $DCFH_2$-DA assay, with an assessment based on the variance between technical replicates, fitting of the standard curve based on fluorescein fluorescence, and time- or concentration-related saturation of DCF signal. Data presented for this study used fluorescence values taken at 0 min subtracted from those at 90 min, given that a linear relationship was observed during this time period.

In addition to using the fluorescein fluorescence to provide a molar value for $DCFH_2$ oxidation, the data from each partner was normalized based on a comparison to the positive (NPCB and SIN-1) controls to provide a percentage change based on these high values, or to negative (medium only) controls as a ratio of change. Linear regression of log-transformed values with post-hoc pairwise comparisons was

Table 1. The key conditions selected for the provisional SOP.

Parameter	SOP condition
NM dispersant	MEM + 2% FCS
Test NM concentration range	1.6, 3.1, 6.3, 12.5, 25, 50, 100 μg/ml
Particle positive control	NPCB: 1.6, 3.1, 6.3, 12.5 μg/ml
Chemical positive control	SIN-1: 1.3, 2.5, 5, 10 μM
$DCFH_2$-DA concentration	10 μM
Time points	30, 60, and 90 min
F-DA standard curve	0.001, 0.004, 0.012, 0.037, 0.111, 0.333, 1.000 μM

performed to compare both the dose-response slopes and mean values across the different partners utilizing each of the different normalization methods. Dose-response curves using AU were further assessed to allow the effective concentration (EC) 50 calculation by fitting a Hill-slope. Comparison of data achieved in intra-laboratory comparison was assessed by ANOVA and Tukey post-hoc analysis. Statistical analysis was completed using Stata (v16) or SPSS.

Results

Optimization of initial SOP parameters

Before the provisional SOP was written a number of parameters were tested to ensure that optimum conditions were selected. These are described in more detail in the supplemental material, and therefore only briefly summarized here. The selected conditions for the key parameters of the initial SOP are summarized in Table 1.

Two commonly applied chemical positive controls, tert-Butyl hydroperoxide (t-BHP) and SIN-1, were compared (Figure S1). SIN-1 was shown to be far more reliable than t-BHP, in terms of the linear relationship between dose and signal, the reproducibility across replicates, and the stability of these oxidants overtime after preparation. SIN-1 was therefore suggested as a more reliable positive control within the SOP.

Fluorescence generated by two different $DCFH_2$ concentrations (10 and 50 μM) were tested over time in response to various concentrations of the chemical positive control, SIN-1 (Figure S2). Although absolute fluorescence signals were similar for each $DCFH_2$ concentration, when presented as a ratio to the negative control, the higher background signal of 50 μM resulted in less sensitive signals in response to SIN-1 and therefore 10 μM $DCFH_2$ was selected for the SOP.

The impact of using water or cell culture medium (MEM) in the preparation of NMs for the assay, as well as the effect of including fetal calf serum (FCS) on the fluorescence signal, was assessed using NPCB, CuO, and titanium dioxide NMs (TiO_2 NMs) (Figure S3). The use of MEM as a suspension media resulted in a lower signal compared to water, which was further reduced by the inclusion of FCS for each NM tested. However, as shown in Table S1, the use of FCS provided a far better particle dispersion than without FCS, identified by a lower particle diameter in all cases, and a lower polydispersity index (PDI). The improvement in dispersion induced by FCS was true for both suspensions in water and in MEM. Although water with FCS demonstrated a slightly better dispersion than MEM with FCS, the presence of salts and protein in the MEM suspension media is considered more reflective of the physiological environment of human exposure, therefore for the remaining studies, NMs were prepared in MEM containing 2% FCS.

The dispersion method was further optimized using various dispersion protocols as described in Table S2. Briefly, methods of alcohol pre-wetting, suspension using water with FCS, and suspension immediately in assay-relevant media were compared, alongside a comparison of preparation vessels, using both plastic and glass. The most reliable particle suspensions were obtained using an initial dispersion of particles in plastic bijous tubes at 1 mg/ml in phenol red-free MEM plus 2% FCS, followed by sonication for 15 min in a sonicating water bath; these conditions performed better than when different dispersion aids were tested, and better than when suspension as in glass or in other plastic vessels such as Eppendorf tubes.

The stability of F-DA with various storage conditions and the reproducibility of performing multiple batch preparations (Figures S4 and S5), where it was shown that samples maintained a consistent and repeatable dose-dependent fluorescence signal after storage at −20 °C. This was true for the same batch assessed at different time points during a 12-week period, for different batches prepared and stored for periods of up to 8 weeks, and for different batches prepared and assessed on the same day. Increasing F-DA concentrations were included on each experimental assay plate to generate a standard curve for fluorescein production.

The impact of plate location bias or cross-well contamination of a fluorescence signal was tested under the conditions of this SOP (Figures S6 and S7), and as none were found, it was possible to preserve identical plate formats throughout testing.

The use of F-DA for identifying fluorescence quenching (to control for potential assay interference by NM) is shown in Figure S8. The optimal concentration for detecting an interference effect was determined to be 0.1 μM.

Interlaboratory comparison – ranking of representative materials

Five contributing partners measured $DCFH_2$ oxidation in response to NMs of the following materials ZnO, $BaSO_4$, Mn_2O_3, and CuO, following the first iteration of the SOP (Figure 1, Table 2 and supplemental material). Although the absolute fluorescence values were often considerably different across partners, all partners identified ZnO and $BaSO_4$ as low-ROS-producing and Mn_2O_3 and CuO as high-ROS-producing NFs. Potential interference of the NMs in the assay was tested, with respect to either fluorescence quenching and enhancement or auto-fluorescence, and none was found (Figures S9 and S10).

First interlaboratory comparison and effectiveness of normalization

Results generated by each partner were compared over time and NM concentration, to an NM positive control (NPCB), a chemical positive control (SIN-1), a standard curve of

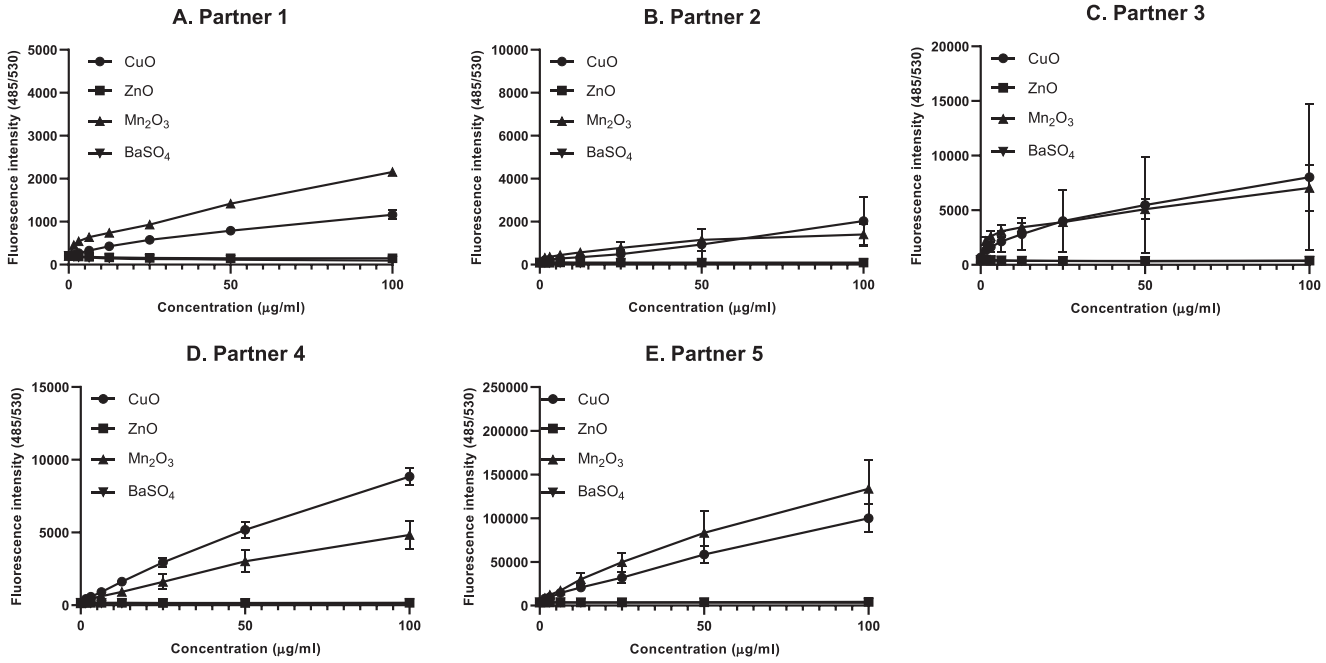

Figure 1. (A–E) Interlaboratory comparison by five partners of $DCFH_2$ oxidation by ZnO, BaSO4, Mn_2O_3, and CuO using provisional SOP (conducted during first interlaboratory comparison). Data represents the mean and SD of at least three independent replicates, and are expressed as arbitrary fluorescence units.

Table 2. Hydrodynamic diameter (by DLS) of ZnO, BaSO4, Mn_2O_3 and CuO when suspended in MEM with 2% FCS after sonication in different ultra-sonicating water baths for 15 min.

	Partner 1		Partner 2		Partner 3		Partner 4	
Sonication bath	Ultrawave, QS25		Ultrawave, QS25		Bandelin Sonorex RK510		Elmasonic S100	
Energy input (J/s)	400		400		160		150	
Sample	Z-av. (nm)	PDI	Z-av. (nm)	PDI	Z-av. (nm)	PDI	Z-av. (nm)	PDI
CuO	178.2 ± 27.4	0.4 ± 0.06	285.5 ± 118.6	0.53 ± 0.09	275.6 ± 206.8	0.55 ± 0.18		
Mn_2O_3	279.8 ± 6.8	0.23 ± 0.01	362.4 ± 27.9	0.34 ± 0.06	297.1 ± 87.3	0.39 ± 0.15	348.6 ± 11.4	0.25 ± 0.01
ZnO	462.7 ± 33.9	0.33 ± 0.03	452.4 ± 149.5	0.42 ± 0.11	496.1 ± 75.6	0.42 ± 0.07	380.2 ± 10.2	0.2 ± 0.01
BaSO4	283.6 ± 46.5	0.6 ± 0.09	281.5 ± 134.3	0.6 ± 0.11	549.1 ± 305.7	0.59 ± 0.11	300.7 ± 10.2	0.25 ± 0.15
NPCB	338.5 ± 19.3	0.52 ± 0.08	285.9 ± 56.1	0.49 ± 0.11	408.8 ± 112.6	0.67 ± 0.13	281 ± 8.1	0.33 ± 0.03

Data shows average particle diameter (Z.av) and uniformity of these measurements with a polydispersity index (PDI).

fluorescein fluorescence, and negative control of medium only. The 90 min time point was chosen for most comparisons, although some consideration has been given to earlier time points. Statistical analysis was performed using linear regression, and dose-response curve evaluation.

$DCFH_2$ oxidation by SIN-1, NPCB, CuO, and Mn_2O_3 measured by all partners in the first interlaboratory comparison study were plotted side-by-side in Figure 2. The data are expressed as arbitrary fluorescence units as well as with each of the normalization methods applied, that is, by standard curve, positive controls, and negative control. ZnO and $BaSO_4$ were excluded from further data analysis as these NM did not generate a response greater than background values. For the first interlaboratory comparison study, there was a clear variation between partners in AU with orders-of-magnitude differences in the concentration-response curve found between partners for each of the NM and SIN-1 control. Normalization of data to F-DA standard curve, positive controls (NPCB and SIN-1), and negative control (medium only) appeared to reduce the magnitude of variation observed between partners.

Using the pairwise comparisons, it was possible to identify a maximum number of statistically significant differences that were possible when comparing different partners. For any and every given data set (i.e. each of the individual graphs in Figure 2), there were ten pairwise comparisons made in the first interlaboratory study: partner 1 to partners 2, 3, 4, and 5, partner 2 to partners 3, 4, and 5, and so forth. However, due to only four partners conducting the second interlaboratory study, only six pairwise comparisons were possible. To allow direct comparison of the performance of the two interlaboratory studies we, therefore, have expressed these pairwise comparisons as a percentage of total possible, that is, out of ten in the first study, and out of six in the second study. Two metrics were included in each pairwise comparison, the differences in mean fluorescence AU and the differences in the slope of the exposure concentration-response, represented in Figure 3 by whole and dashed lines, respectively.

Prior to data normalization, there were substantial differences found between partners' mean data values, with almost all possible data comparisons between partners indicating significant differences. It was observed that 70%, 90%,

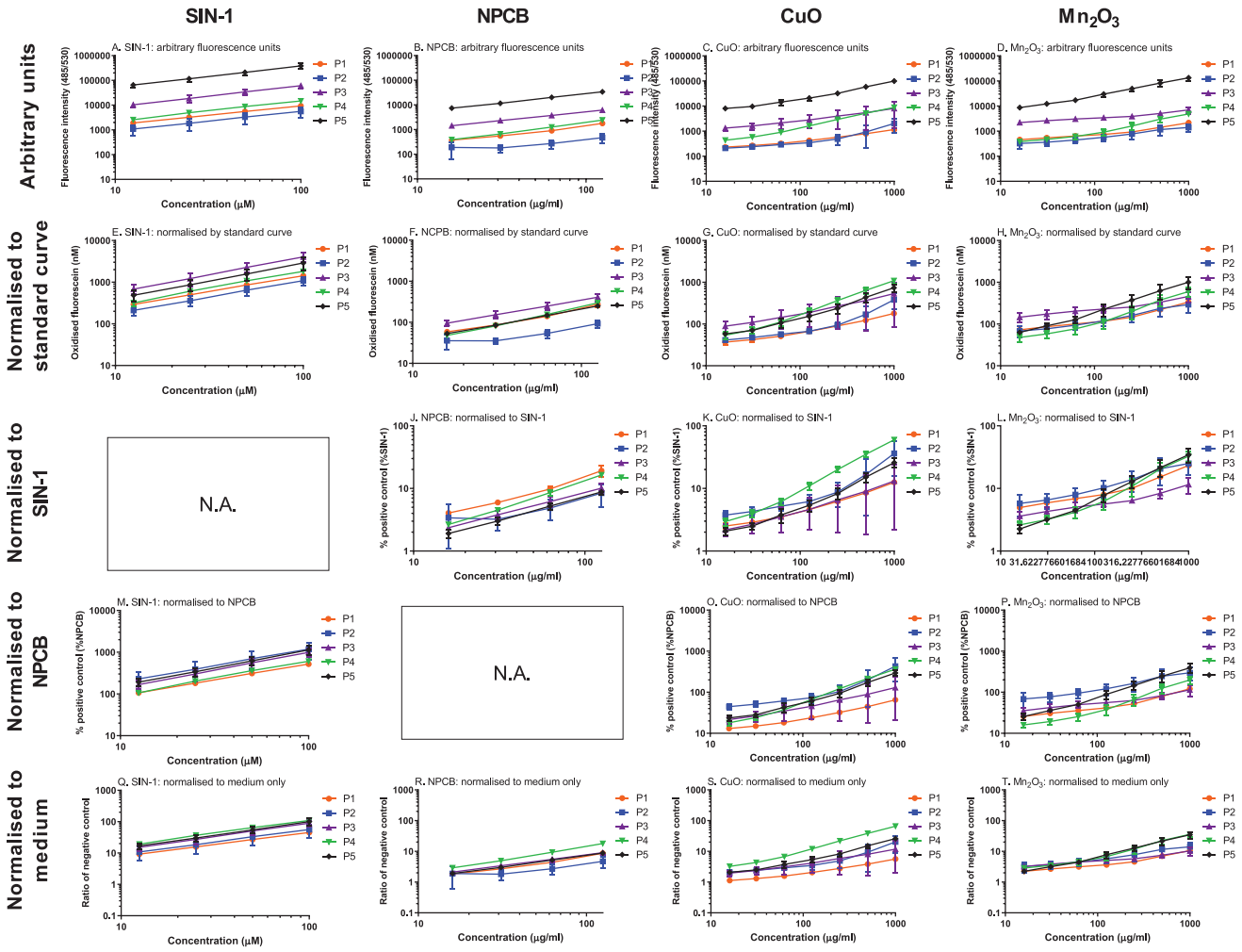

Figure 2. First interlaboratory comparison, conducted with the original SOP, of $DCFH_2$ oxidation by SIN-1, NPCB, CuO, and Mn_2O_3, and expressed as arbitrary fluorescence units as well as by each method of normalization: standard curve, positive controls (SIN-1, NPCB) and negative control (medium). The data is of five contributing partners, is presented as a log-scale for both axis, and represents the mean and SD of at least three independent replicates. (N.A. Not applicable).

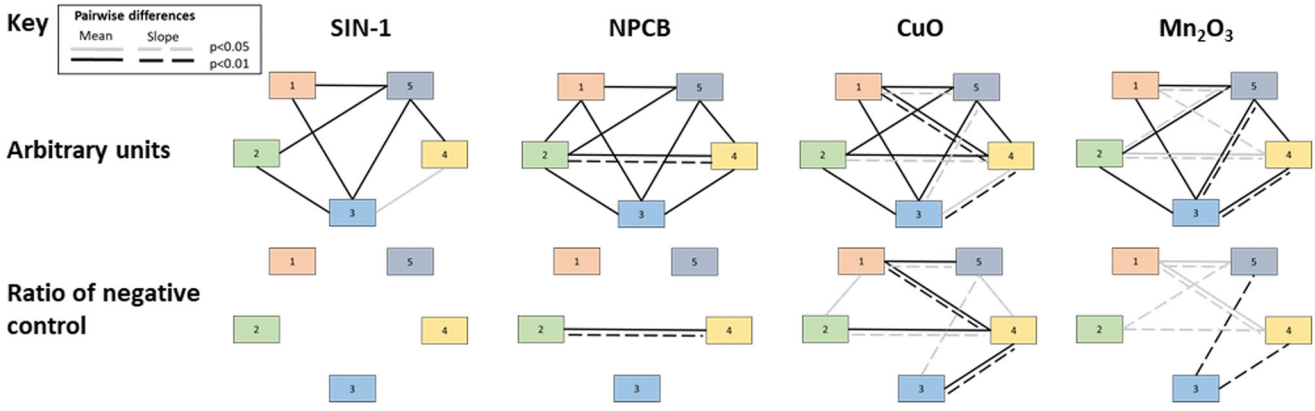

Figure 3. Pairwise differences between partners during first interlaboratory comparison study. In each column are differences shown in evaluation of SIN-1, NPCB, CuO, or Mn_2O_3, and rows are differences found when using arbitrary fluorescence units or when data was normalized to negative control values; statistically significant differences between mean values are shown with lines, and between slopes with dotted lines.

90%, and 80% of all pairwise comparisons of mean fluorescence AU were significantly different for SIN-1, NPCB, CuO, and Mn_2O_3 data, respectively. All methods of normalization (extrapolation from standard curve, and comparison to positive and negative controls) reduced the statistically significant differences between partner slopes and means, indicating a more reproducible data set; however, the most striking change was the normalization using the SIN-1 data to standardize results of other substances, where differences in mean fluorescence among all partners were no longer statistically significant after normalization. When comparing the slopes of the exposure concentration-response curves,

Table 3. Number of pairwise differences between partners in interlaboratory comparison studies.

1st interlaboratory study						2nd interlaboratory study					
Number of pairwise differences observed between participating laboratories											
Comparison of means (significant differences: out of a potential 10)						Comparison of means (significant differences: out of a potential 6)					
	Arb. units	Mol.	% SIN-1	% NPCB	Med.		Arb. units	Mol.	% SIN-1	% NPCB	Med.
SIN-1	7	1		0	0	SIN-1	2	1		0	0
NPCB	9	5	2		1	NPCB	3	3	1		0
CuO	9	6	4	4	6	CuO	3	3	4	1	1
Mn_2O_3	8	5	1	5	2	Mn_2O_3	2	0	3	3	3
Comparison of slopes (significant differences: out of a potential 10)						Comparison of slopes (significant differences: out of a potential 6)					
	Arb. units	Mol.	% SIN-1	% NPCB	Med.		Arb. units	Mol.	% SIN-1	% NPCB	Med.
SIN-1	0	0		0	0	SIN-1	0	0		0	0
NPCB	1	1	1		1	NPCB	0	0	0		0
CuO	5	5	5	5	5	CuO	0	0	0	0	0
Mn_2O_3	6	6	6	6	6	Mn_2O_3	3	3	3	3	3
Percentage of total possible pairwise differences – using data presented in upper half of table											
Comparison of means (significant differences: % of maximum)						Comparison of means (significant differences: % of maximum)					
	Arb. units	Mol.	% SIN-1	% NPCB	Med.		Arb. units	Mol.	% SIN-1	% NPCB	Med.
SIN-1	70	10		0	0	SIN-1	33	17		0	0
NPCB	90	50	20		10	NPCB	50	50	17		0
CuO	90	60	40	40	60	CuO	50	50	67	17	17
Mn_2O_3	80	50	10	50	20	Mn_2O_3	33	0	50	50	50
Comparison of slopes (significant differences: % of maximum)						Comparison of slopes (significant differences: % of maximum)					
	Arb. units	Mol.	% SIN-1	% NPCB	Med.		Arb. units	Mol.	% SIN-1	% NPCB	Med.
SIN-1	0	0		0	0	SIN-1	0	0		0	0
NPCB	10	10	10		10	NPCB	0	0	0		0
CuO	50	50	50	50	50	CuO	0	0	0	0	0
Mn_2O_3	60	60	60	60	60	Mn_2O_3	50	50	50	50	50

For each individual data set (identified as individual graphs in Figures 2 and 4) the number statistically significant differences between the data collected by participating laboratories were calculated; in this assessment a high number reflects a low level of reproducibility across participating laboratories. Significant differences were noted when $p < 0.05$ for mean values as well as exposure concentration-response curve slopes, and expressed as a number per total possible differences, or as % of total possible differences. The columns noted as Arb. Units, Mol., % SIN-1, % NPCB and Med. represent statistical analysis of data from raw fluorescence arbitrary units (Arb. Units), or of data analyzed after normalization by extrapolation of a molar fluorescein concentration using the standard curve (Mol.), or normalization through comparison to positive (% SIN-1, % NPCB) and negative (Med.) controls.

SIN-1 and NPCB were already similar across partners, with either 0 or 10% differences found, respectively, which did not change with normalization; CuO and Mn_2O_3 slope data were more dissimilar for AU (50%) and after normalization, with 60% of possible differences found, regardless of normalization method applied. An example of where these statistically significant differences within the first interlaboratory comparison can be seen is given in Figure 3 for AU and normalization by negative control, with pairwise data (*p*-values and confidence intervals for this example shown in Tables S3 and S4); the complete number of statistically significant pairwise differences are shown in Table 3, with all other comparisons shown in Figure S11.

SOP refinement and second interlaboratory comparison

Following analysis of the initial data set, the SOP was refined and interlaboratory comparison was repeated by 4 partners for SIN-1, NPCB, CuO, and Mn_2O_3 (Figure 4). Refinement to the SOP was minimal as the key conditions outlined in Table 1 remained unchanged, however, a more detailed SOP was generated to better align the assay process, for example, timings for reagent preparation, selection of plastics. The final SOP is fully described in SI.

The use of a revised SOP in the second interlaboratory comparison appeared to improve the reproducibility between partners; for partners' mean fluorescence AU 33%, 50%, 50%, and 33% of possible significant differences were found for SIN-1, NPCB, CuO, and Mn_2O_3 data, respectively. Normalization of SIN-1 data, by extrapolation from the standard curve and comparison to positive and negative controls, was shown to improve this reproducibility, while reproducibility of the NPCB data was improved with normalization by comparison to the negative control or SIN-1. Normalization of CuO and Mn_2O_3 data by any method did not improve reproducibility. The reproducibility of the slopes obtained for SIN-1, NPCB, and CuO was greatly improved in the second interlaboratory comparison compared to the first, with no significant differences observed between partners. In contrast reproducibility of the slope for Mn_2O_3 data between labs was only partially improved (from 60% differences between partners in the first interlaboratory study compared to 50% differences in the second study), and not changed with data normalization. All the specific pairwise comparisons for the second interlaboratory comparison can be seen in Figure S12.

The reproducibility between partners was further assessed through the generation of EC50 values (Figure 5 and Table 4). The SIN-1 dose-response was reproduced well across all

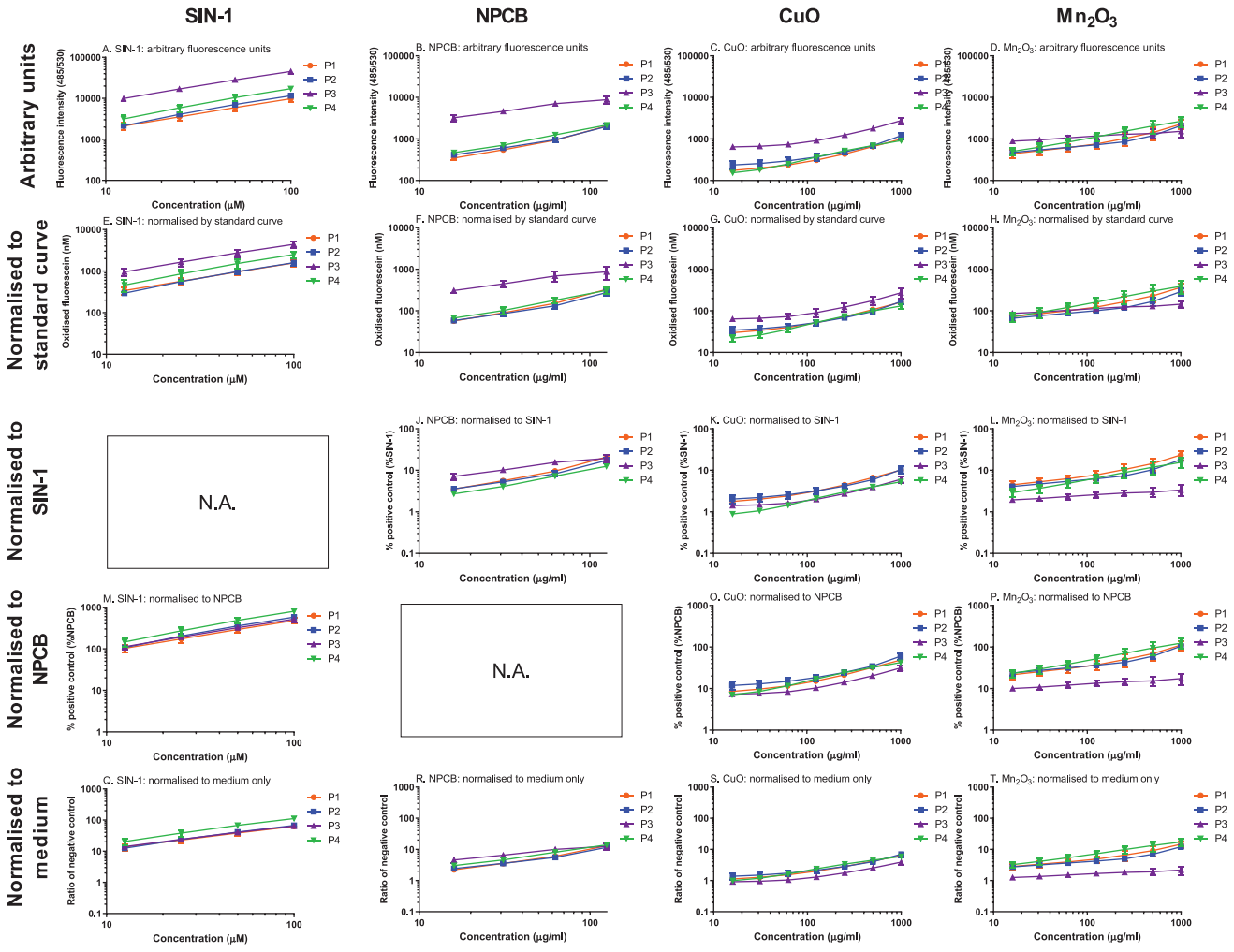

Figure 4. (A–T) Second interlaboratory comparison, conducted with the revised SOP, of DCFH2 oxidation by SIN-1, NPCB, CuO and Mn2O3, and expressed as arbitrary fluorescence units as well as by each method of normalization: standard curve, positive controls and negative control. The data is of four contributing partners, is presented as a log-scale for both axis, and represents the mean and SD of at least three independent replicates.

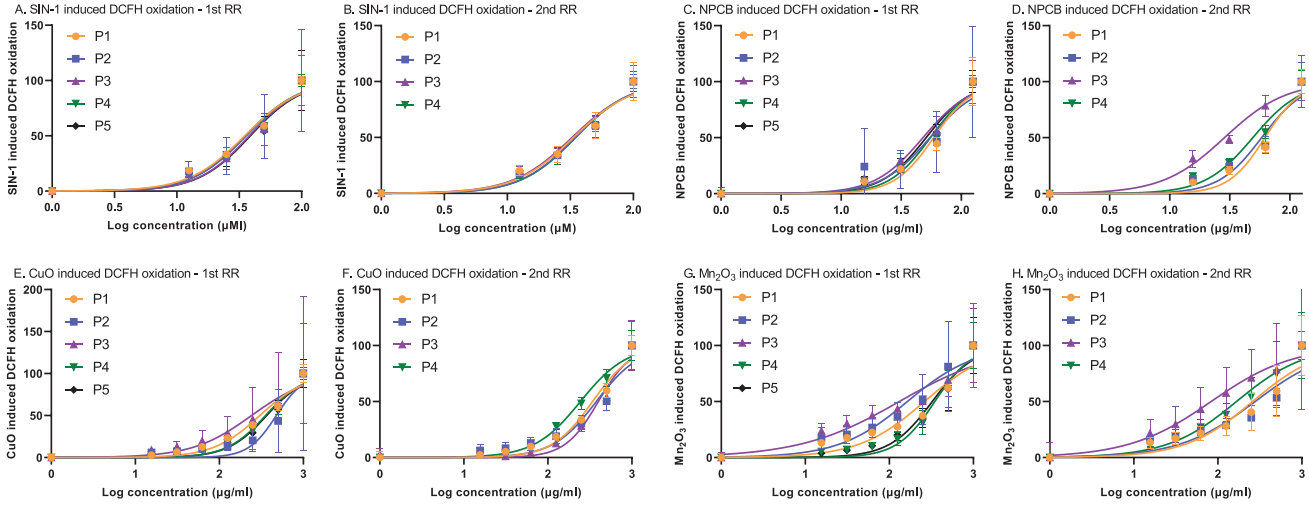

Figure 5. Dose response curve (as hill-slope) for EC50 calculations of data collected in the original (A, C, E, G) and revised (B, D, F, H) SOP for $DCFH_2$ oxidation by SIN-1 (A and B), NPCB (C and D), CuO (E and F) and Mn_2O_3 (G and H). The data was collected by four contributing partners, and represents % effect over log transformed particle concentrations, and is the mean and SD of at least three independent replicates; $DCFH_2$ oxidation was recorded over a 90 min period.

Table 4. EC50 values for $DCFH_2$ oxidation by SIN-1, NPCB, CuO and Mn_2O_3 collected during the original and revised SOP.

	P1	P2	P3	P4	P5	Av.	SD	COV
	1st interlaboratory comparison – data from 30 min							
SIN-1	37.0	35.9	37.0	33.4	40.0	36.6	2.4	6.5
NPCB	62.9	68.1	52.1	57.1	54.6	59.0	6.5	11.0
CuO	360.9	559.2	229.5	430.7	429.2	401.9	120.1	29.9
Mn_2O_3	252.2	154.5	93.2	380.8	310.3	238.2	115.9	48.7
	1st interlaboratory comparison – data from 60 min							
SIN-1	35.3	35.6	37.9	34.7	39.0	36.5	1.9	5.1
NPCB	55.5	61.6	49.8	55.0	51.7	54.7	4.5	8.2
CuO	331.2	527.8	257.8	388.0	397.7	380.5	99.4	26.1
Mn_2O_3	223.9	133.5	121.5	363.7	332.3	235.0	111.1	47.3
	1st interlaboratory comparison – data from 90 min							
SIN-1	35.0	36.0	38.1	35.6	38.8	36.7	1.6	4.5
NPCB	58.9	54.8	48.0	55.3	50.9	53.6	4.2	7.9
CuO	303.2	477.1	241.6	354.8	365.8	348.5	87.1	25.0
Mn_2O_3	262.3	164.0	131.5	347.9	324.5	246.0	95.7	38.9
	2nd interlaboratory comparison – data from 90 min							
	P1	P2	P3	P4	–	Av.	SD	COV
SIN-1	33.5	34.5	32.7	34.9	–	33.9	1.0	3.0
NPCB	62.0	58.2	28.6	49.2	–	49.5	15.0	30.2
CuO	340.0	377.0	398.7	244.1	–	340.0	68.3	20.1
Mn_2O_3	258.7	277.4	83.1	166.6	–	196.5	89.7	45.7

$DCFH_2$ oxidation was recorded over a 30, 60, or 90 min period.

partners in both studies, with an average EC50 of 36.7 μM, with 1.6 SD and a coefficient of variation (COV) of 4.5 for the original SOP and an average concentration of 33.9 μM, with 1.0 SD and a COV of 3.0 COV for the revised SOP. Data collected for NPCB by the original SOP was also relatively consistent across partners, with an average EC50 of 53.6 μg/ml with 4.2 SD and 7.9 COV. However this was not reproduced with the revised SOP, although the average EC50 was similar at 49.5 μg/ml, the variation between partners was greater, with 15.0 SD and 30.2 COV. EC50 values for CuO and Mn_2O_3 were similar across each study for individual partners, but the large variation seen between partners in the first interlaboratory comparison was not improved with the revised SOP Table 4.

The EC50 values from dose curves generated at different time points of incubation of NMs with $DCFH_2$ were also compared. In general, the COV was greater at earlier time points compared to later time points (Table 4). The full range in EC50 values can be seen in Table S5. The reproducibility of each partner was further assessed through examination of variation within partner replicates and can be seen in Table S6. As a generalization, one partner was shown to be relatively consistent across both studies, while three partners improved the consistency of replicates with the use of the revised SOP, reflected by the changes in COV observed for analysis of SIN-1, NPCB, CuO, and Mn_2O_3. This is an important distinction, as the revised SOP incorporated the methodology of this first partner, with other partners adopting this methodology.

Intralaboratory comparison

The results from two contributing partners (Partner 1 and Partner 2) operating within the same laboratory facility were compared. Each partner followed each iteration of the SOP, and the results were compared based on AU only (Figure 6). The 90 min time point was chosen for these comparisons, and statistical analysis was performed using ANOVA.

Using the provisional SOP a significant difference in ROS generation was detected between the contributing partners at the highest concentrations of SIN-1 (100 μM), NPCB (62.5 and 125 μg/ml), and Mn_2O_3 (1000 μg/ml). However with optimization of the SOP, the same treatment panel was assessed (Figure 7), and no significant differences were observed.

Discussion

In general, the interlaboratory comparison demonstrated that arbitrary fluorescence units show high levels of partner variability, and that normalization improved variability but did not remove it. Excluding the NMs which did not show ROS production in this assay (ZnO and $BaSO_4$), $DCFH_2$ oxidation by all test substances (Mn_2O_3, CuO, and NPCB) were shown to be significantly variable across the contributing partners. However, this is true only when the magnitude of the effect was considered, and this effect was improved upon by introducing methods of normalizing the collected data, with the greatest improvement evident in the use of the SIN-1 control, or the normalization to the negative/background control values. Reproducibility was further improved with optimization of the SOP, with less significant differences observed between contributing partners when considering the magnitude of effect, again further improved with data normalization, particularly to the values of the negative controls.

Performance of the SOP

There are a number of factors that can be considered in the design of a reliable protocol, in fact, there has previously been a cause-and-effect assay design process reported for the development of *in vitro* tests, initially for the MTS assay (Rösslein et al. 2015). In this process, the authors noted a number of sources of assay variability, including cell maintenance, pipetting, instrument performance, chemical positive control, assay protocol, and engineered NM handling and characterization. In terms of *instrument performance*, we incorporated a number of control and calibration treatments to assess whether the use of different instruments, with different performances, could still provide reliably similar results. We also provided an analysis template that contained certain validation checks to ensure that the data collected by each partner did not present any discordant results as a result of instrument limitations. We tested, and confirmed the reliability of chemical positive control (SIN-1), as well also included a positive particle control (NPCB); the *assay protocol*, contained many of the considerations identified and discussed by Rösslein et al. (2015), including age and storage time of reagents, a method to account for background signal levels, and timings to account for optical degradation of reagents; for engineered NM *handling*, an equivalent dispersion protocol was used, and interference within the assay was tested. In their recent publication, Petersen et al. (2020) identify specific parameters relating to cause-and-effect assay design for the $DCFH_2$-DA assay in cells, including instability

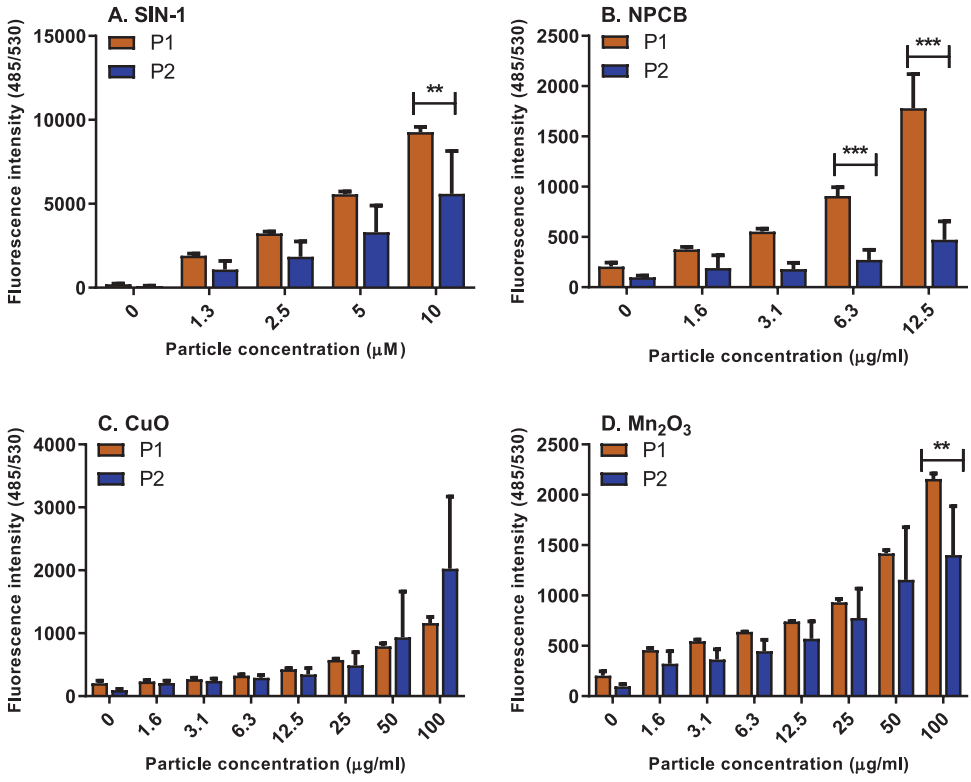

Figure 6. Intralaboratory comparison, conducted with the original SOP, of $DCFH_2$ oxidation by (A) SIN-1, (B) NPCB, (C) CuO, and (D) Mn_2O_3, and expressed as arbitrary fluorescence units. The data represents those of two contributing partners operating from the same facility, and is presented as the mean and SD of at least three independent replicates. Statistically significant differences are shown with $^{**}p < 0.01$ and $^{***}p < 0.005$.

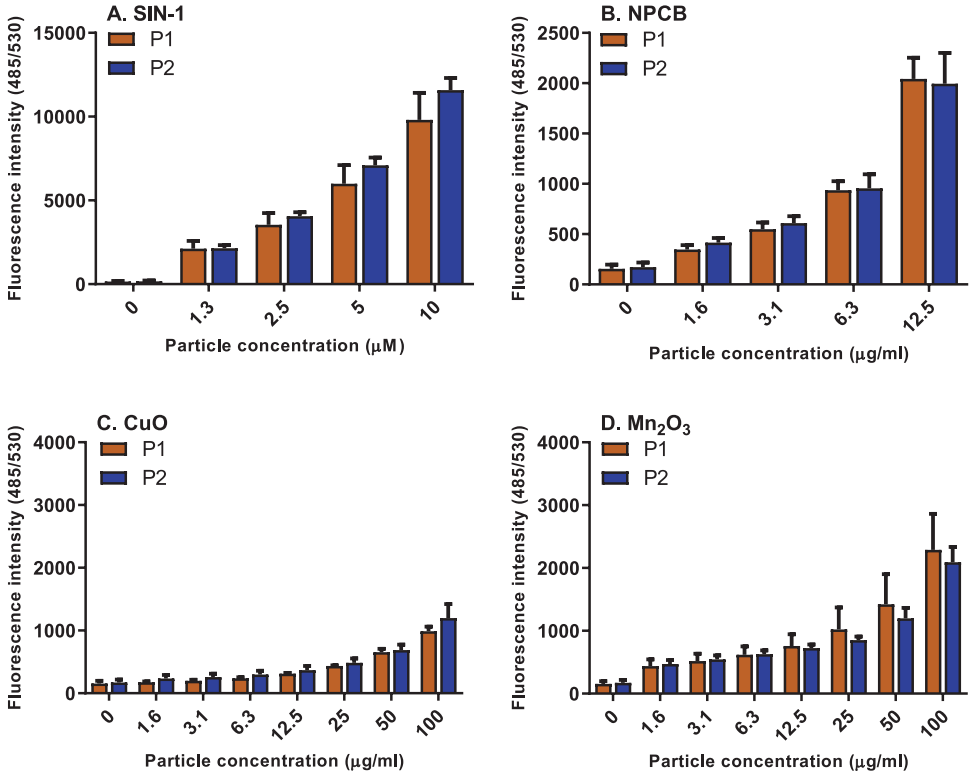

Figure 7. Intralaboratory comparison, conducted with the revised SOP, of $DCFH_2$ oxidation by (A) SIN-1, (B) NPCB, (C) CuO, and (D) Mn_2O_3, and expressed as arbitrary fluorescence units. The data represents those of two contributing partners operating from the same facility, and is presented as the mean and SD of at least three independent replicates.

of the $DCFH_2$-DA reagent with time, and the potential of it to be degraded by light. Our assay design would already address these issues with efforts to protect reagents from light, as well as operate under a relatively strictly timed procedure.

In the assessment of an exposure concentration-response it was already evident in the first interlaboratory comparison that for some substances (particularly NPCB) the level of reproducibility was already very high; with the statistical comparisons made between contributing partners for this data set alone there would be a possible ten significant differences (Partner (P)1 to P2, P3, P4, and P5; P2 to P3, P4, and P5; P3 to P4 and P5; P4 to P5), and only one significant difference out of this possible 10 was observed. CuO and Mn_2O_3 both were shown to present a better reproducibility when considering the shape of the exposure concentration-response curve rather than the magnitude of effect. With the revised SOP, the exposure concentration-response comparison was greatly improved, with no significant differences observed between partners for NPCB or CuO NMs, while Mn_2O_3 was slightly improved. The strength in reproducibility demonstrated by NPCB under this metric was further demonstrated by the closeness of the EC50 concentrations. In general, for the NMs used, we would suggest that measurement of ROS produced by NPCB was attained with the greatest level of accuracy, followed by CuO and then Mn_2O_3; moreover, the chemical control used, SIN-1, was demonstrated as particularly reproducible.

We have opted for use of a 90 min period for use in data analysis, but this was reliant on the fluorescence signal maintaining a linear increase to ensure that no saturation of the reaction had occurred (a factor addressed in our data analysis template). The use of shorter times has been discouraged by Pal et al. (2012) as being unreliable for use in ranking of NM intrinsic ROS-generating potential. Our data presented here is in line with the suggestion by Pal et al. (2012). COV between EC50 values obtained by the contributing partners increased with decreasing incubation time, meaning that the reproducibility of data collected by different laboratories was improved by extending the measurement time to 90 min.

Refinement of the SOP resulted in a more detailed and definitive protocol and included the addition of a loading plate for particle preparation. The success of this revised protocol was promising, but not completely conclusive. When using EC50 values as a guide, an improved reproducibility was shown for SIN-1 and CuO NMs, while ROS generated by NPCB and Mn_2O_3 were found more consistent with the original protocol. However, when considering the statistically significant differences in the magnitude and slope, assessed by linear regression, there were less significant differences between contributing partners in all test materials in the second protocol compared to the first. It is possible that the better-defined particle handling and the more precise reagent preparation and timing did have an impact here, and improved reproducibility, which may be expected considering the points raised by Rösslein et al. (2015) and Petersen et al. (2020) for sources of variation in assays such as this one.

Intralaboratory comparison

Although the first round of interlaboratory testing had achieved promising results, the main driver in designing and performing the second round of testing using an optimized SOP was the significant variability for SIN-1, NPCB, and Mn_2O_3 found between the results of two partners using the same laboratory equipment. The second round of testing with further assay optimization improved this intra-laboratory reproducibility considerably, with operatives demonstrating ROS generation with no statistically significant differences for all substances tested. It is not clear exactly why the differences were present initially, the main changes during the SOP development were selecting specific plastics to be used, providing a harmonized dispersion protocol, and providing a more prescribed sequence for reagent preparation. The known oxidation of $DCFH_2$ in the air (Doak et al. 2009) lends support to the use of a clearly defined preparation process, whereby the conditions and times are aligned and potential sources of variation reduced (Rösslein et al. 2015; Petersen et al. 2020). This should improve reproducibility, and here this was the case. Furthermore, there have been reported concerns surrounding pipetting-error-derived variation (Rösslein et al. 2015), the use of certain plastics causing inaccuracies in particle concentrations (Holzwarth et al. 2019), and an inadequate level of protocol detail (Piret et al. 2017), all relevant as possible sources of variation and resulting in a lack of reproducibility.

Data analysis – recommended normalization and assay controls

We had selected NPCB as a positive particle control, given its well-established high level of ROS generation. As mentioned above, NPCB was demonstrated to be considerably reliable within both interlaboratory comparison studies, particularly when data normalization was used, and especially when the dose-response curves were compared. As for our chemical control, we had initially compared SIN-1 and t-BHP for reliability; with a strong dose-response and a stable signal over time and over different storage conditions, we recommend SIN-1 as a good assay control. Rösslein et al. (2013) argue that due to the spontaneous generation of ROS by SIN-1 it is only suitable as a qualitative positive control in a cellular assay. However, here in an acellular format, we provide evidence of the stability of SIN-1 not only over long storage conditions, and in the assessment of batch-to-batch variability, but also with an extended time period left on the bench before an assay is completed. SIN-1 was consistently shown to be reliable across different replicates, users, and laboratories, with almost identical EC50 values obtained by five different partners in two separate studies; this aligns with the expectations of a suitable chemical positive control stipulated by Rösslein and colleagues in their recent publication (Petersen et al. 2020). The stability reported here supports the choice of SIN-1 over t-BHP as chemical control, as SIN-1 is easier to handle and generates more consistent results.

In terms of how well these controls provided a robust means of normalization, SIN-1 improved reproducibility of mean values from 90% dissimilarity between partners measuring ROS formed by NPCB in the first interlaboratory comparisons (using arbitrary units) to just 20% dissimilarity, with CuO improved from 90% dissimilarity to 40% dissimilarity, and Mn_2O_3 from 80% to 10%. With the use of the signal generated by NPCB to normalize other data sets, signals of arbitrary units obtained for SIN-1, CuO and Mn_2O_3 were improved from 70%, 90%, and 80% dissimilarity, to 0%, 40%, and 50% dissimilarity respectively. However, the greatest improvement was consistently provided by the use of the medium-only negative controls, with SIN-1, NPCB, CuO and Mn_2O_3 detected using arbitrary units improving reproducibility from 70%, 90%, 90%, and 80% dissimilarity, respectively, to 0%, 10%, 60%, and 20%, respectively.

In general, in the second interlaboratory study, only the normalization to the medium control was most successful in improving reproducibility, with SIN-1, NPCB, CuO and Mn_2O_3 detected using arbitrary units improving reproducibility from 33%, 50%, 50%, and 33% dissimilarity, respectively, to 0%, 0%, 17%, and 50%, respectively; the reproducibility of Mn_2O_3 was not improved by normalization.

Data analysis – use of F-DA standard calibration curve and in interference assays

As a means to determine if any test material would cause interference by fluorescence quenching or enhancement, F-DA was proposed. We also wished to test the potential to use increasing concentrations of F-DA as a calibration curve to allow the formation of DCF to be converted from fluorescence arbitrary units to relative fluorescein molarity; similar approaches for calibration of the DCF signal have been used, including with a reference to an H_2O_2 calibration curve (Pal et al. 2012). In order to support the use of F-DA for either of these applications we first tested the reproducibility of F-DA batch preparation and stability of F-DA over long periods of cold storage. We also determined an F-DA concentration which provided the optimal sensitivity in the detection of fluorescent quenching, as well as the F-DA concentration range which generates the optimal calibration curve covering the fluorescence range of the $DCFH_2$-DA assay.

The use of F-DA as a calibration curve and extrapolation of a molar value for $DCFH_2$ oxidation was least effective of all normalization methods (excluding Mn_2O_3 detection in the second interlaboratory study, in which dissimilarity between partners was reduced to 0%, however, this was not consistent). Therefore, we do not recommend this method for this purpose. However, F-DA is still useful to test interference by fluorescent quenching. The assessment of interference is critical. There has previously been demonstrated an inverse dose-response for some materials tested by the $DCFH_2$-DA assay (Pal et al. 2012). It is possible that this is due to undetected interference with the fluorescence signal, and this concern is relevant for both cellular and acellular forms of the assay (Doak et al. 2009; Kroll et al. 2012; Aranda et al. 2013). However, it can be accounted for with adequate testing of interference. Here we propose F-DA for this purpose, while elsewhere it is reported to be done with $DCFH_2$ itself (Doak et al. 2009). Here we have accounted for this interference and restricted which particle concentrations to test based on this information. By doing so we have observed strong dose-response data for each of the materials which had demonstrated an ability to oxidize $DCFH_2$.

Recommendations for further optimization of SOP

Although the range of ROS detected by $DCFH_2$ has not been fully resolved, and there are dissimilarities and similarities with other assays reported earlier which demonstrate the requirement to consider this factor when using this method; using cellular assays it has previously been shown that the ROS produced during respiratory burst is detected sufficiently by $DCFH_2$, and that $DCFH_2$ actually responds well to nitrogen species, to $RO_2^{\cdot}$, $RO^{\cdot}$, $OH^{\cdot}$, HOCl and ONOO\bar, not directly to $O_2^{\cdot}$\bar, and only responds well to H_2O_2 in the presence of peroxidase activity (Walrand et al. 2003; Foucaud et al. 2007; Doak et al. 2009). It, therefore, is possible to include peroxidase enzymes, as has previously been done for the assessment of NPCB (Foucaud et al. 2007) and of carbon nanotubes (Rothen-Rutishauser et al. 2010), but it should be acknowledged that up to 30% of $DCFH_2$ oxidation may be in response to the enzyme itself (Walrand et al. 2003), which can increase the background signal within the assay (Pal et al. 2012); making the addition of HRP difficult to support.

The impact on particle dispersion is likely to impact on reproducibility; here, although contributing partners used the same sonication method, the water bath sonication apparatus were not aligned and therefore the sonication power applied to the test materials would differ and range from 150 to 400 J/s (Table 2); this may have affected the level of particle dispersion. Size distributions were measured by DLS, and can be seen as different, but also each was also shown to have considerably large PDI values, indicating a heterogeneous mix of particle agglomerates. The impact of the dispersion method has been shown elsewhere to impact particle-generated ROS detected by $DCFH_2$ (Pal et al. 2012).

Oxidative potential of representative materials

Grouping and read-across are approach methodologies well defined for chemicals and under development for nanoforms (NFs) (Giusti et al. 2019). As already outlined (Stone et al. 2020), this rationale is advocated for use within a framework used for establishing reliable grouping approaches of NFs; in which an Integrated Approach to Testing and Assessment (IATA) of NF hazard can include varying levels (acellular/*in chemico*, *in vitro* and *in vivo*), each with varying complexity and increased confidence. One basis of the grouping approach is an NF's physicochemical properties, including reactivity and therefore a measure of NF surface reactivity and generation of ROS. This is based on the accepted paradigm that intrinsic ROS production by some NFs can drive toxicity and represents a mechanism by which NFs cause cellular toxicity via oxidative stress. Here the $DCFH_2$-DA assay was able to distinguish between NMs which produce high

levels of ROS and those that produce low, or no ROS. However, it was not sensitive enough to distinguish differences between low ROS-producing NM (e.g. ZnO) and those producing no ROS ($BaSO_4$) (Ag Seleci et al. 2022). This was similar across all contributing partners. Equally, a high level of ROS production was detected by all partners for both CuO NMs and Mn_2O_3 NMs, as well as the positive control NPCB. These results, in part, reflect findings elsewhere, including a high level of oxidative potential demonstrated for CuO NMs and Mn_2O_3 NMs in a cytochrome C assay (Delaval et al. 2017), the ferric reducing ability of serum (FRAS) assay, and in electron paramagnetic resonance (EPR) measurements (Ag Seleci et al. 2022). Conversely, when using the cytochrome C assay (Delaval et al. 2017), FRAS and EPR with DMPO spin trap (Ag Seleci et al. 2022) a distinction, albeit a small one, has been made between $BaSO_4$ and ZnO NMs, when here we have no such distinction, and neither was one found when using EPR with a tempone-h spin trap. Interestingly, it is specifically superoxide ($O_2^{\cdot}$\bar) which is measured with tempone-h, and amongst limitations to the $DCFH_2$-DA assay, numerous accounts of its limited ability to be oxidized in the presence of H_2O_2, but also $O_2^{\cdot}$\bar (Doak et al. 2009), and potentially it is this specificity that result in there being a lack in sensitivity of the $DCFH_2$-DA assay to certain substances, including ZnO. Incidentally, when reported to react to $O_2^{\cdot}$\bar, it is thought likely that $DCFH_2$ is responding to peroxynitrite (ONOO\bar) formed in the presence of $O_2^{\cdot}$\bar (Walrand et al. 2003), and as noted by Rösslein et al. (2013), $DCFH_2$ is also used for detection of hydroxyl radicals ($HO^{\cdot}$).

Conclusions

This SOP for the $DCFH_2$-DA assay has been used to demonstrate a robust intra-laboratory reproduction of ROS measurements from numerous NMs, while efficient reproduction across different laboratories was also obtained, although extent of reproducibility was particle-type-specific.

Acknowledgments

The technical assistance from Angus Hargan (HWU) and Yasmin Akhtar (NRCWE) was greatly appreciated.

Disclosure statement

W.W. is an employee of BASF SE, a company producing and marketing nanomaterials. All other authors declare no conflict of interest.

Funding

This work was financially supported by the European Commission-funded project GRACIOUS [760840].

ORCID

Matthew Boyles http://orcid.org/0000-0003-4291-4903
Nicklas Raun Jacobsen http://orcid.org/0000-0002-2504-2229

References

Ag Seleci D, Tsiliki G, Werle K, Elam DA, Okpowe O, Seidel K, Bi X, Westerhoff P, Innes E, Boyles M, Miller M, Giusti A, Murphy F, Haase A, Stone V, Wohlleben W. 2022. Determining nanoform similarity via assessment of surface reactivity by abiotic and In vitro assays. NanoImpact. Submitted.

Aranda A, Sequedo L, Tolosa L, Quintas G, Burello E, Castell JV, Gombau L. 2013. Dichloro-dihydro-fluorescein diacetate (DCFH-DA) assay: a quantitative method for oxidative stress assessment of nanoparticle-treated cells. Toxicol in Vitro. 27(2):954–963.

Bass DA, Parce JW, Dechatelet LR, Szejda P, Seeds MC, Thomas M. 1983. Flow cytometric studies of oxidative product formation by neutrophils: a graded response to membrane stimulation. J Immunol. 130: 1910.

Bilski P, Belanger AG, Chignell CF. 2002. Photosensitized oxidation of 2′,7′-dichlorofluorescin: singlet oxygen does not contribute to the formation of fluorescent oxidation product 2′,7′-dichlorofluorescein. Free Radic Biol Med. 33(7):938–946.

Brandt R, Keston AS. 1965. Synthesis of diacetyldichlorofluorescin: a stable reagent for fluorometric analysis. Anal Biochem. 11:6–9.

Daghastanli NA, Itri R, Baptista MS. 2008. Singlet Oxygen Reacts with 2′,7′-dichlorodihydrofluorescein and contributes to the formation of 2′,7′-dichlorofluorescein. Photochem Photobiol. 84(5):1238–1243.

Delaval M, Wohlleben W, Landsiedel R, Baeza-Squiban A, Boland S. 2017. Assessment of the oxidative potential of nanoparticles by the cytochrome c assay: assay improvement and development of a high-throughput method to predict the toxicity of nanoparticles. Arch Toxicol. 91(1):163–177.

Doak SH, Griffiths SM, Manshian B, Singh N, Williams PM, Brown AP, Jenkins GJS. 2009. Confounding experimental considerations in nanogenotoxicology. Mutagenesis. 24(4):285–293.

EC. 2006. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC.

ECHA. 2016. New approach methodologies in regulatory science. Paper presented at a Scientific Workshop; April 19–20; Helsinki, Finland.

Foucaud L, Wilson MR, Brown DM, Stone V. 2007. Measurement of reactive species production by nanoparticles prepared in biologically relevant media. Toxicol Lett. 174(1–3):1–9.

Giusti A, Atluri R, Tsekovska R, Gajewicz A, Apostolova MD, Battistelli CL, Bleeker EAJ, Bossa C, Bouillard J, Dusinska M, et al. 2019. Nanomaterial grouping: existing approaches and future recommendations. NanoImpact. 16:100182.

Hellack B, Nickel C, Albrecht C, Kuhlbusch TAJ, Boland S, Baeza-Squiban A, Wohlleben W, Schins RPF. 2017. Analytical methods to assess the oxidative potential of nanoparticles: a review. Environ Sci: Nano. 4(10):1920–1934.

Holzwarth U, Cossío U, Llop J, Kreyling WG. 2019. Unpredictable nanoparticle retention in commonly used plastic syringes introduces dosage uncertainties that may compromise the accuracy of nanomedicine and nanotoxicology studies. Front Pharmacol. 10:1293.

Karlsson HL, Cronholm P, Gustafsson J, Möller L. 2008. Copper oxide nanoparticles are highly toxic: a comparison between metal oxide nanoparticles and carbon nanotubes. Chem Res Toxicol. 21(9): 1726–1732.

Kroll A, Pillukat MH, Hahn D, Schnekenburger J. 2012. Interference of engineered nanoparticles with in vitro toxicity assays. Arch Toxicol. 86(7):1123–1136.

Pal AK, Bello D, Budhlall B, Rogers E, Milton DK. 2012. Screening for oxidative stress elicited by engineered nanomaterials: evaluation of acellular DCFH assay. Dose-Response. 10(3):308–330.

Pang C, Neubauer N, Boyles M, Brown D, Kanase N, Hristozov D, Fernandes T, Stone V, Wohlleben W, Marcomini A. 2017. Releases from transparent blue automobile coatings containing nanoscale

copper phthalocyanine and their effects on J774 A1 macrophages. NanoImpact. 7:75–83.

Petersen EJ, Hirsch C, Elliott JT, Krug HF, Aengenheister L, Arif AT, Bogni A, Kinsner-Ovaskainen A, May S, Walser T, et al. 2020. Cause-and-effect analysis as a tool to improve the reproducibility of nanobioassays: four case studies. Chem Res Toxicol. 33(5):1039–1054.

Piret J-P, Bondarenko OM, Boyles MSP, Himly M, Ribeiro AR, Benetti F, Smal C, Lima B, Potthoff A, Simion M, et al. 2017. Pan-European inter-laboratory studies on a panel of in vitro cytotoxicity and pro-inflammation assays for nanoparticles. Arch Toxicol. 91(6):2315–2330.

Rösslein M, Elliott JT, Salit M, Petersen EJ, Hirsch C, Krug HF, Wick P. 2015. Use of cause-and-effect analysis to design a high-quality nanocytotoxicology assay. Chem Res Toxicol. 28(1):21–30.

Rösslein M, Hirsch C, Kaiser J-P, Krug HF, Wick P. 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(12): 24320–24337.

Rothen-Rutishauser B, Brown DM, Piallier-Boyles M, Kinloch IA, Windle AH, Gehr P, Stone V. 2010. Relating the physicochemical characteristics and dispersion of multiwalled carbon nanotubes in different suspension media to their oxidative reactivity in vitro and inflammation in vivo. Nanotoxicology. 4(3):331–342.

Sarvazyan N. 1996. Visualization of doxorubicin-induced oxidative stress in isolated cardiac myocytes. Am J Physiol. 271(5):H2079–H2085.

Sauvain J-J, Rossi MJ, Riediker M. 2013. Comparison of three acellular tests for assessing the oxidation potential of nanomaterials. Aerosol Sci Technol. 47(2):218–227.

Schlinkert P, Casals E, Boyles M, Tischler U, Hornig E, Tran N, Zhao J, Himly M, Riediker M, Oostingh GJ, et al. 2015. The oxidative potential of differently charged silver and gold nanoparticles on three human lung epithelial cell types. J Nanobiotechnology. 13:1.

Stone V, Gottardo S, Bleeker EAJ, Braakhuis H, Dekkers S, Fernandes T, Haase A, Hunt N, Hristozov D, Jantunen P, et al. 2020. A framework for grouping and read-across of nanomaterials- supporting innovation and risk assessment. Nano Today. 35:100941.

Walrand S, Valeix S, Rodriguez C, Ligot P, Chassagne J, Vasson M-P. 2003. Flow cytometry study of polymorphonuclear neutrophil oxidative burst: a comparison of three fluorescent probes. Clin Chim Acta. 331(1–2):103–110.

Wang H, Joseph JA. 1999. Quantifying cellular oxidative stress by dichlorofluorescein assay using microplate reader11Mention of a trade name, proprietary product, or specific equipment does not constitute a guarantee by the United States Department of Agriculture and does not imply its approval to the exclusion of other products that may be suitable. Free Radical Biol Med. 27(5–6):612–616.

Wilson MR, Foucaud L, Barlow PG, Hutchison GR, Sales J, Simpson RJ, Stone V. 2007. Nanoparticle interactions with zinc and iron: implications for toxicology and inflammation. Toxicol Appl Pharmacol. 225(1):80–89.

Wilson MR, Lightbody JH, Donaldson K, Sales J, Stone V. 2002. Interactions between ultrafine particles and transition metals in vivo and in vitro. Toxicol Appl Pharmacol. 184(3):172–179.