Determining what really counts: Modeling and measuring nanoparticle number concentrations

Elijah J. Petersen[1], Antonio R. Montoro Bustos[2], Blaza Toman[3], Monique Johnson[2], Mark Ellefson[4], George C. Caceres[2], Anna Lena Neuer[6], Qilin Chan[5], Jonathan Kemling[5], Brian Mader[4], Karen Murphy[2], Matthias Roesslein[6]

[1] Biosystems and Biomaterials Division, Material Measurement Laboratory, National Institute of Standards and Technology (NIST), 100 Bureau Drive, Gaithersburg, MD 20899

[2] Chemical Sciences Division, Material Measurement Laboratory, National Institute of Standards and Technology (NIST), 100 Bureau Drive, Gaithersburg, MD 20899

[3] Statistical Engineering Division, Information Technology Laboratory, National Institute of Standards and Technology (NIST), 100 Bureau Drive, Gaithersburg, MD 20899

[4] 3M, Environmental Laboratory, St. Paul, MN, USA

[5] 3M, Corporate Research Analytical Division, St. Paul, MN, USA

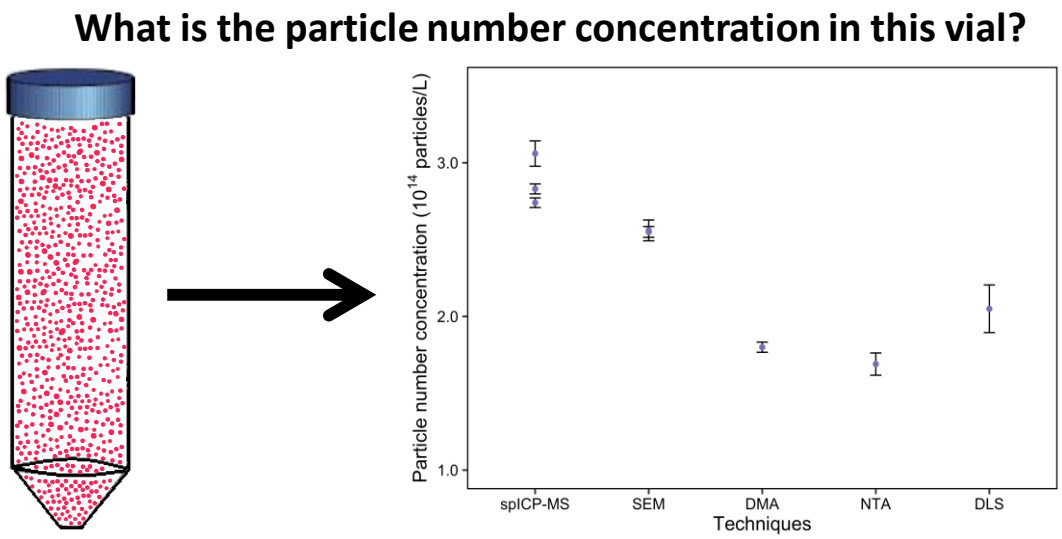
[6] EMPA, Swiss Federal Laboratories for Material Testing and Research, Particles-Biology Interactions Laboratory, CH-9014 St. Gallen, Switzerland

Abstract

Particle number concentration (PNC) measurements are critical for research and regulatory decision making related to the potential applications and implications of nanotechnology. However, the degree to which different analytical methods yield similar PNCs has not yet been studied. In this study, monodisperse gold nanoparticles (AuNPs) with varying sizes (30 nm or 60 nm) and surface coatings (citrate, polyvinylpyrrolidone, or branched polyethyleneimine) were evaluated using five techniques: scanning electron microscopy (SEM), dynamic light scattering (DLS), differential mobility analysis (DMA), nanoparticle tracking analysis (NTA), and single particle inductively coupled plasma-mass spectrometry (spICP-MS). The two techniques that only measured the NP core size (spICP-MS and SEM), as opposed to the larger hydrodynamic diameter, yielded PNCs with the closest agreement (within 20 % of each other), while PNCs among all techniques sometimes varied by a factor of 3. Positively charged AuNPs coated with branched polyethyleneimine yielded the most variable results. Deriving the PNC using the particle size distribution has several advantages over using only the mean size based on these results and statistical modeling given the substantial impact of the tails of the distribution toward smaller particles. The size distributions measured by the different techniques were also used to model the AuNP concentration that would reach the cells in an *in vitro* toxicity experiment. Surprisingly, there was a strong impact of the analytical technique on the modeled cellular AuNP concentration for some of the AuNPs.

Keywords: Nanotechnology, particle counting, nanotoxicology, size distribution

TOC graphic



This paper describes a comprehensive investigation of particle number concentrations including a multi-method comparison, theoretical modeling, and cellular dosimetry.

Introduction

The enhanced or novel properties of nanoparticles (NPs) are expected to lead to their widespread use in consumer products such as in polymeric materials,[1-3] and for commercial applications in textiles, biomedical applications, and environmental applications.[1-5] NPs are defined as particles with one dimension between 1 nm and 100 nm.[6, 7] During the life cycle of these materials, it is possible that NPs will be released causing exposure to workers, consumers, and ecological receptors.[4, 8-12] This has led to extensive research to develop robust methods to assess potential toxicological risks[13-21] and to quantify NPs in different matrices (e.g., water, soil, and biological tissues).[22-28]

The issue concerning which dose metric to use in nanotoxicological studies (i.e., whether to use the mass, particle number, or surface area concentration to assess the response of cells or organisms to NP exposure) has been a topic of debate since nearly the beginning of the nanotoxicology field.[29] While measuring a mass concentration for dissolved organic and inorganic substances is linearly related to their number concentration, the situation is more complex for NPs. Unlike dissolved chemicals, NPs have a distribution of sizes; they may undergo changes in test media such as dissolution or agglomeration; and the conversion from a mass- or surface area-based concentration to a number-based concentration requires a more complex formula than a simple linear correlation. Although the mass concentration is the most widely reported metric for the exposure concentration in nanoecotoxicological research, some studies have suggested that alternative dose metrics such as the surface area-based or a particle number concentration (PNC)-based metric may more accurately reflect the toxicological response observed.[30-33] In addition, recent research efforts have also been made to evaluate the size distribution of NPs associated with test organisms after exposure and the PNC of the organisms' body burden after extraction from the tissue and resuspension into a liquid media.[15, 34-37]

In addition to their importance in nanotoxicology, PNC measurements also have regulatory importance.[38] For example, the use of a PNC as the metric in some geographical locations such as the European Union (e.g., 50 % of the particles between 1 nm and 100 nm) has been proposed for determining if a substance is labelled as containing NPs.[39] In addition, one key consideration for the use of OECD ecotoxicology test guidelines with NPs is what dose metric to use when evaluating if the change in the exposure concentration has exceeded the limit of ± 20 %.[40]

One of the principle challenges in determining the PNC of NPs suspended in aqueous media using size distribution measurements is that different analytical procedures can give varying results. This stems partly from the potential for even a small number (1 %) of NP agglomerates to shift the whole size distribution to larger particle sizes for some techniques such as nanoparticle tracking analysis (NTA) and dynamic light scattering (DLS), given that these techniques are much more sensitive to larger particles. Other techniques such as single particle inductively coupled plasma-mass spectrometry (spICP-MS) would count agglomerates as part of the tail toward larger particles, but this would be unlikely to shift the full size distribution.[41] In addition, NP size measurement techniques also measure slightly different properties of the NPs with some measuring only the NP core diameter (e.g., scanning electron microscopy (SEM) and spICP-MS), while other techniques (e.g., DLS and NTA) also measure the hydrodynamic diameter, which

includes the NP surface coating (if the NP is being stabilized) and hydrated water ions. Therefore, previous studies typically show that DLS and NTA results yield larger diameters for NPs than results from other techniques that only measure the NP core.[41-43] When converting from the measured size to the NP number concentration, it is unclear to what degree these size differences would impact the PNC. For some techniques that directly measure the NP core, there are also limitations such as low throughput analysis and challenges with sample deposition for SEM analysis.[44] Some techniques also directly measure the NP number concentration such as spICP-MS, NTA, and potentially differential mobility analysis (DMA).[45-47] However, it is unclear to what degree these direct PNC measurements agree among techniques or with PNC values derived using NP size measurements.

There have been several studies that have compared PNC measurements across laboratories for the same initial NPs using a single technique. Among studies utilizing spICP-MS, one study was performed by a *post hoc* analysis of previously published spICP-MS data for the National Institute of Standards and Technology (NIST) reference material (RM) 30 nm and 60 nm gold nanoparticles (AuNPs),[45] and two other studies assessed silver nanoparticles (AgNPs) in food simulants[48] or after addition to chicken meat.[49] An interlaboratory comparison has also been conducted on polystyrene NPs and 30 nm AuNPs using NTA.[47] In the spICP-MS interlaboratory comparison of AuNP results, the PNC recoveries for the 60 nm AuNP ranged between 63.9% and 99.95%, while the PNC recoveries for the 30 nm AuNP ranged between 14.8% and 102.2%, suggesting that larger NPs may yield better recoveries.[45] Results for AgNPs yielded an even broader range with the average recovery (after removal of outliers) ranging between 0.6% and 39% compared to the expected values from the manufacturer.[48] This result that could stem from numerous factors including particle dissolution, losses from adsorption to the containers,[50] and how the transport efficiency was calculated.[51] However, there has not yet been a comparison among techniques for measuring PNCs.

In this study, we conducted a multi-technique (Table 1) and multi-laboratory study to investigate the comparability of PNC results for four AuNPs. To minimize variability that could result from NP dissolution, matrix effects from complex aqueous matrices, or agglomeration as a result of a high ionic strength media, a simple scenario was evaluated, namely AuNPs in water. Four monodisperse AuNPs were tested: two NIST RMs (8012 and 8013) and two commercially available AuNPs with different surface coatings which impacted the surface charge (positively-charged branched polyethyleneimine (bPEI) and negatively-charged polyvinylpyrrolidone (PVP)). Three samples were negatively charged: citrate-stabilized, AuNPs NIST RM 8012 and RM 8013 with nominal diameters of 30 nm and 60 nm, respectively, and the PVP AuNPs, while the bPEI AuNPs were positively charged (Table S1). Because the measured mean values and shape of the size distributions were found to vary among techniques, statistical analysis was performed to understand the impact of variations in these and other parameters on the derived PNC results. The size distributions measured by the different techniques were also used to model the AuNP concentration that would reach the cells in an *in vitro* toxicity experiment, an approach that has been used to evaluate the toxicological effects of NPs on, for example, human macrophage[52] and alveolar epithelial cells.[53]

Methods

Test materials

Four aqueous dispersions of different monodisperse AuNPs with approximate spherical shapes were tested in this study. 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. For the RMs, particle sizes were previously extensively characterized at NIST with detailed information provided in the ROIs.[54, 55] The other two samples, purchased commercially, were PVP and bPEI coated AuNP suspensions with nominal diameters of 30 nm. The characteristics of the AuNPs suspensions studied here, are provided either in the NIST ROI or by the manufacturer are given in Table S1. The identity for the two NIST RMs was revealed to the analysts at all three laboratories while the other two samples were unknown (except for one spICP-MS analyst and the total Au analysts at one laboratory who were aware of the properties for all four NPs).

NIST RM 8012 and NIST RM 8013 aqueous suspensions were supplied in 5 mL hermetically sealed pre-scored glass ampoules sterilized by gamma irradiation. For both commercial AuNP samples, 5 mL aliquots were supplied in glass vials or in Nalgene bottles. Ice packs were used to keep the samples at 4 °C during shipping.

Laboratory 1: Total gold analysis

The mass fraction of Au in the test materials at various timepoints was quantified by ICP-MS throughout the study. The purpose of these measurements is that they were used to derive the PNC values. For laboratory 1, two to three nominal, 0.25 g subsamples per vial were accurately weighed into individual, clean, low density polyethylene (LDPE) bottles. The mass of each sub-sample was recorded to ± 0.00001 g. Following this, 0.1 mL of concentrated nitric acid (HNO_3) and 0.3 mL of concentrated hydrochloric acid (HCl) (both Optima grade, Thermo Fisher Scientific, Waltham, MA, USA) were added and the samples allowed to digest at room temperature for 15 h. Sample solutions were observed to turn from pink to colorless. Process blanks, composed of 0.25 g of water, were treated in the same manner as samples. All water used for sample processing was prepared in-house by sub-boiling distillation using a conditioned, quartz still with deionized water as feedstock. An accurately weighed mass of platinum (Pt) internal standard solution, prepared from NIST SRM 3140 Platinum Standard Solution, was added to each sample and process blank (collectively referred to as "samples"). Samples were then quantitatively diluted with 10 g of water, forming the first serial dilution. Samples were quantitatively diluted a second time using an aqueous diluent solution composed of 0.5 % thiourea (w/v), 2.4 % HCl (v/v) and 0.4 % HNO_3 (v/v).

Mass spectrometric analyses were performed on a ThermoFisher Scientific X series II ICP-MS equipped with matrix tolerant (Xt) cones and operated at 1400 W. Solutions were introduced *via* a peristaltic pump into a low-flow (100 µL/min) PFA micro-concentric nebulizer. The nebulizer was fitted to an impact-bead spray chamber cooled to 2 °C. Measurements were made in continuous mode using peak jump data acquisition with one point per peak. Three to five blocks of data, each one minute in duration, were acquired per sample. Signal intensities at m/z 195 and 197 were recorded. Duplicate mass spectrometric analyses were acquired per sample. The mass

fraction of Au in each sample was computed using an external calibration curve. Au standards spanning the range from 0.5 µg/kg Au to 14 µg/kg Au prepared gravimetrically from NIST SRM 3121 Gold Standard Solution in the same thiourea/acid diluent solution as the samples, were used to construct the calibration curve. Temporal changes in signal intensity throughout the mass spectrometric analyses were corrected *via* the Pt internal standard present at similar mass fraction in all samples and standards.

The mass fraction of Au measured in this manner includes contributions from Au present both as AuNPs and as ionic Au. In order to assess whether any free ionic Au (i.e. not bound as AuNPs) was present in the test materials, nominal 0.2 g subsamples were accurately weighed into 15 mL centrifuge tubes, followed by the addition of nominal 5 mL water. Samples were then centrifuged at – 15 °C for 1 h at 7000 *g* (Allegra 25R fixed angle rotor, Beckman Coulter). Two, nominal 1 mL subsamples of the supernatant were carefully withdrawn from the centrifuged samples and accurately weighed into LDPE bottles. A known mass of Pt internal standard was added, the samples were diluted gravimetrically in 4 mL of the thiourea/acid diluent, and the mass fraction of Au was measured in the manner described above.

Laboratory 2: Total gold analysis

All AuNP suspensions were prepared at four different dilutions with Milli-Q® 18.2 MΩ·cm ultrapure water. Triplicate aliquots (0.5 mL) of each dilution level were transferred to 15 mL polyethylene screw-capped tubes (Corning Sciences, Corning, New York) and a total of 12 replicates were analyzed for each AuNP. One mL of freshly prepared aqua regia ((3:1 v/v, HCl (BDH VWR Analytical, Radnor, PA): HNO_3 (Honeywell Fluka, Mexico City, Mexico)) was added to each tube and put onto a mixer for 30 min. Once dissolved, the solutions were diluted to a final volume of 10 mL with 2 % L-cysteine hydrochloride monohydrate (Sigma Aldrich, St. Louis, MO) (w/v) in Milli-Q® 18.2 MΩ·cm ultrapure water. The samples were diluted an additional five-fold with 2 % HCl (v/v), 2 % L-cysteine hydrochloride monohydrate (w/v) in Milli-Q® 18.2 MΩ·cm ultrapure water.

Calibration standards for a four-point calibration curve were prepared by diluting the elemental Au standard purchased from High Purity Standards (Charleston, SC) with 2 % HCl (v/v), 2 % L-cysteine hydrochloride monohydrate (w/v) in Milli-Q® 18.2 MΩ·cm ultrapure water. Aliquots of three samples were selected as laboratory control spikes (LCS). An elemental gold standard solution was spiked into the LCS samples at three different concentrations: 0.2 mg/L, 0.4 mg/L, and 0.6 mg/L.

The instrument used for the elemental analysis was a Perkin Elmer Optima 8300DV ICP optical emission spectrophotometer (Waltham, PA). The samples were analyzed against a four-point external calibration curve. A rinse solution containing 2 % HCl (v/v) and 2 % L-cysteine hydrochloride monohydrate (w/v) was used to minimize instrument carry-over between samples. The rinse time was set to 180 s. A 0.5 mg/L quality control standard was used to assess the accuracy of the calibration curve during the analysis. A 0.5 mg/L solution of scandium (High Purity Standards) was run in-line with the samples and standards to serve as an internal standard.

Laboratory 1: Single particle inductively coupled plasma-mass spectrometry (spICP-MS) analysis

Single particle ICP-MS measurements of all samples in Laboratory 1 were conducted using a Thermo Electron X Series X7 quadrupole ICP-MS quadrupole ICP-MS system (Waltham, MA, USA) with a C-type nebulizer (0.5 mL/min) and an impact bead spray chamber cooled to 2 °C. Descriptions of the spICP-MS technique and all other analytical techniques are provided in Table 1. The instrument was tuned daily to a minimum $^{156}CeO/^{140}Ce$ oxide level (<2%) and a maximum ^{115}In sensitivity. The sample flow rate was set to approximately 0.45 mL/min, and the uptake rate was measured daily, in triplicate, by weighing the water uptake after 5 min of aspiration. NIST RM AuNP suspensions were prepared in triplicate by serial dilution of stock suspensions with deionized water to an approximate particle number concentration of 15000 particles/mL. However, for the two unknown samples (PVP and bPEI AuNPs), three different dilution levels were tested to obtain an adequate PNC that provided a sufficient number of events for counting statistics and that minimized the particle coincidence occurrence. A blank (deionized water) and at least five soluble Au standards ranging from 0 to 100 ng/g Au in a thiourea solution (0.1 % thiourea (w/v), 2.4 % HCl (v/v), and 0.5 % HNO_3 (v/v)) were measured to calculate the Au sensitivity of the instrument. NIST RM 8013 was used as the NP calibrant for all materials except the RM 8013 measurements, in which case, RM 8012 was used. Since an AuNP standard was used, the measurement of Au standards was not necessary, but this was performed in order to assess differences in transport efficiencies computed by the frequency and size methods, a finding recently reported in a thorough study on this topic.[51] As differences were observed, the transport efficiency calculated via the frequency method was used for direct PNC quantification, whereas a response factor (expressed in counts per second per ng of Au) established from signal intensities measured for RM 8013 was used to measure the particle size distribution (PSD).[56-58] For spICP-MS measurements of AuNPs, the signal for ^{197}Au was recorded using time-resolved analysis mode with Thermo Fisher PlasmaLab software using a 10 ms dwell time. Data were exported to Microsoft Excel for data processing. Ionic standard solutions were analyzed for 180 s, while AuNP standards and suspensions were measured three times for 360 s for a total of 1080 s. A threshold of particle intensities five standard deviations above the mean signal intensity was chosen as the criteria for distinguishing between single particle events and the signal from dissolved ions in solution. Particle sizes were calculated for all single particle events.

Laboratory 2: Single particle inductively coupled plasma-mass spectrometry (spICP-MS) analysis

Single particle ICP-MS measurements of all samples in Laboratory 2 were conducted using an Agilent Technologies, Inc. 7900 ICP-MS system (Santa Clara, Ca) with a MicroMist nebulizer and a Scott-style double-pass spray chamber. The instrument was auto-tuned daily. The sample flow rate was set to deliver 0.346 mL/min. All AuNP suspensions were prepared in triplicate by serial dilution of stock suspensions with 18.2 MΩ·cm ultrapure water to an approximate particle number concentration of 15000 particles/mL. The instrument was calibrated using an ionic blank (1 % HCl (v/v)) and a soluble Au standard of 1 ng/g Au in 1 % HCl (v/v). For spICP-MS measurements of AuNPs, RM 8013 was used as the NP calibrant for all samples except the RM 8013 measurements, in which case RM 8012 was used. The signal for ^{197}Au was recorded using

single particle analysis mode with Agilent Technologies MassHunter software (ver. 4.3) using a 0.1 ms dwell time. The MassHunter software calculated the transport efficiency *via* the frequency method for particle number concentration quantification and PSD. The standard solution was analyzed for 120 s, while AuNP suspensions were measured three times for 120 s for a total of 360 s. A threshold of particle intensities five standard deviations above the mean signal intensity was chosen as the criteria for distinguishing between single particle events and the signal from dissolved ions in solution. Data were exported to Microsoft Excel for data processing. Particle sizes were calculated for all single particle events.

Laboratory 3: Nanoparticle tracking analysis (NTA)

All measurements for nanoparticle tracking analysis (NTA) were made using a Malvern NS500Z with software version 3.1. This software incorporates a finite track length adjustment (FTLA) algorithm, that compensates for the size distribution broadening caused by the stochastic nature of the Brownian motion. Hence, all results of NTA provide representative width of the particles, also described as hydrodynamic radii.[59]

Furthermore, the NTA system was calibrated for particle concentration measurements with 100 nm polystyrene NP (Malvern Instruments) at different levels of detector sensitivity. These recordings of the Brownian motion were then analyzed with different amplification settings. This calibration process for particle concentrations allows for accurate detection all particles present in a given observation window for different types of materials and with it different particle surface reflectivity. The number of particles per mL is calculated based on this number providing a direct observation of particle concentration.

The NTA system allows a dynamic or static observation of the investigated particles. The dynamic sample introduction system produces a continuous flow of particles, which allows analysis of between 1500 to 3000 individual particles within 60 s compared with the just 20 to 30 particles in the static arrangement. The dynamic arrangement increases the number of observed particles by a factor close to 100, which of course also improves the statistical robustness of the observed PSD. The dynamic setup was used for all cases, where over the recording period no significant reduction of the particle numbers (which would indicate agglomeration) was observed. If a reduction of the particle number was determined, then any recordings with a significant decreased particle number was excluded from the data analysis. In addition, each time before a repeated series of measurements started, the sample was vortexed rigorously followed immediately by the recording of the Brownian motion. Additional information for specific analysis conditions used to measure the different samples are provided in the SI.

Laboratory 2: Differential mobility analysis (DMA)

A 450 µL aliquot of each well-mixed sample was transferred to a polyethylene micro-centrifuge tube purchased from Axygen (Union City, CA) and capped. The samples were centrifuged in a Beckman Coulter centrifuge (Brea, CA) at 6290 RCF for 12 min. Following centrifugation, 425 µL of supernatant was removed and discarded while the individual nanoparticles remained in a pellet on the bottom of the tube. The pelleted particles were re-suspended by adding 275 µL of 5.0 mmol/L ammonium acetate purchased from JT Baker (Center

Valley, CA) and vortexed for 20 s. The buffer exchange resulted in the re-suspended particles being concentrated by a factor of 1.5. In addition, greater than 90 % of the insoluble sodium citrate buffer was replaced with a volatile buffer which helped to reduce background particle formation during the nebulization process.

The ES/DMA instrument consisted of a TSI 3480 Electrospray Aerosol Generator coupled to a TSI 3082 Electrostatic Classifier and a TSI 3788 Nano Water-Based Condensation Particle Counter (CPC; TSI Incorporated, Shoreview, MN). TSI Aerosol Instrument Manager Software (ver. 10.1.0.6) was used to collect the data. The samples were placed in the pressurized sampling chamber and sprayed through a 0.040 mm diameter capillary. The flow rate of the carrier gas was 1.2 L/min. The dried aerosol then passes to the dynamic mobility analyzer where the voltage is ramped from -12 V to -4.2 kV. The sheath flow in the DMA was set at 15 L/min. The diameter of AuNPs was characterized by electrical mobility, which is inversely proportional to the projected area of the particle. Once sized, the particles travel to the CPC where they were counted. Size distribution data was collected over a minimum of 20 consecutive scans for each unknown sample. Coating AuNPs with insoluble sodium citrate salt during the electrospray process affects the size measurement of particle diameters. Therefore, corrected values for the mobility size of bare AuNPs for NIST 8012, NIST 8013, and PVP AuNP samples were determined using a method described previously[60] using the following equation:

$$d_{p0} = \sqrt[3]{d_{p,m}^3 - d_s^3} \quad (3)$$

We were able to determine a corrected value for the mobility size of bare AuNPs d_{p0}, where $d_{p,m}$ and d_s are mobility sizes measured by DMA of the AuNPs covered with a layer of dried salts and nanoparticles consisting of only the salt itself, respectively.

Laboratory 3: Dynamic light scattering (DLS)

All DLS measurements were performed using a Malvern Nano Zetasizer ZS90. This instrument is equipped with a He-Ne laser 633 nm and it detects the scattered light at a 90-degree angle. The performance of this instrument was periodically evaluated using NIST RM 8012 and 8013. The measurements follow the description of the DLS measurement method given in the NIST ROIs[54, 55]. Briefly, all cuvettes were rinsed and the samples filtered with a 0.1 µm filter (Acrodisc-syringe filter, Pall Corporation) prior to analysis. All measurements were made in the automated mode, where the instrument selected the attenuation factor and then recorded between 11 and 18 runs measuring the dynamic light scattering of the particles. It transformed the variation of the scattered light into the autocorrelation function, from which the instrument selected to 10 best ones for calculating the z-average size and the polydispersity.

NIST RM8013 was diluted by a factor or 10 with MilliQ water (> 18 MΩ·cm), whereas all the other samples were diluted by a factor of 5. Before each measurement the cuvettes were mixed for approximated 10 s using a vortex. Special care was taken to remove any air bubbles, which could have developed during the stirring process.

Laboratory 1: Scanning electron microscopy (SEM)

In Laboratory 1, the NIST ROI size and size distribution values for SEM were used for the RM8012 and 8013 samples, while those for the PVP and bPEI AuNP samples were independently measured. Detailed information for the HR-SEM method for analysis of the PVP and bPEI AuNP samples has been previously described.[58] Briefly, a previously published protocol[61] was used but with a slight modification in that samples were added to Si wafer chips. HR-SEM measurements of clean and individual AuNPs were acquired within 2 d of sample preparation from 6 replicates of each sample and at least 10 individual locations within a selected site on each wafer. Examples of the micrographs obtained for the different AuNPs by the two laboratories are provided in Figure S13.

Laboratory 2: Scanning electron microscopy (SEM)

Samples were prepared for SEM by dilution with ultra-pure electronics-grade water supplied by an in-house water purification unit designed by Smith Engineering (Eden Prairie, MN). Two µL of diluted sample were applied using an adjustable Eppendorf pipettor (Hauppauge, NY) to the surface of a 200 mesh formvar coated copper grid purchased from Ted Pella, Inc. (Redding, CA) to obtain a nominal concentration of approximately 5.0 x 10^6 particles per grid for determination of mean particle diameter by SEM. The samples were allowed to dry and were submitted for imaging without any additional sample preparation.

To ensure statistical significance, a minimum of 200 images were collected for each particle type. A Hitachi SU-8230 Field-Emission Scanning Electron Microscope (Schaumburg, IL) was used to collect the images of the AuNP samples. A scanning transmission electron microscopy (STEM) in a SEM hybrid technique was used which combines through-sample imaging of TEM with the focused rastering electron beam of SEM. The instrument accelerating voltage was set at 30 KeV with a working distance of 8 mm and a tilt of 0 degrees. The bright field-STEM imaging mode was used in STEM mode with a magnification of 100 000 x. Each image was processed with a median filter and a sharpening filter before segmentation.

Image Pro Premier software (Rockville, MD) was used to identify and size particles. For the NIST RM 8012, PVP and bPEI AuNP samples, the default "Dark" segmentation routine in Image Pro Plus was used with smoothing set to 3 and grow set to 1. For the NIST RM 8013 sample, a manual segmentation was applied, selecting all pixels between 0 to 80 on the 0 to 255 grayscale range, with smoothing set to 3 and grow set to 3. Segmented images were analyzed for maximum particle diameter. Images were taken at 100 000 x with an image resolution of 2 560 x 1 920 to ensure that 30 nm diameter particles had ~ 50 pixels across in accordance with NIST procedure PCC-15.[61] The pixel size at these imaging conditions was 0.4961 nm per pixel or 2.016 pixels per nm. The particle diameter values were exported into Excel and each processed image was visually inspected for identification of sizing errors, making sure that the segmentation of particles was correct and did not include multiple particles (doublets, triplets, etc.) or any foreign material that was not an AuNP.

Modeling the cellular concentration

The DG-ISDD model[62] was used to investigate the modelled cellular concentration using the following cellular exposure condition for all measurements: solvent viscosity (0.00081 kg s^{-1}

m^{-1}), solvent density (1.0104 g/cm^3), solvent temperature (37 °C), gold density (19.3 g/cm^3), agglomerate density equivalent to gold density (i.e., no agglomeration), column height (6.0 mm; approximately equivalent to 0.2 mL in a 96-well plate), initial concentration (0.1 mg/cm^3), simulation duration (24 h), no dissolution, and the sticky bottom assumption. The mean size or PSD (using the mass fraction in different size bins) measured using the various techniques for each of the four different AuNPs were also input into the model. It is important to note that changing these parameters, such as the column height, would impact the modeled results. The model was run using Matlab (2017).

Statistical analyses: Particle size distribution (PSD)

For the methods that provide data on each individual particle size (i.e., SEM and spICP-MS), these data were then summarized using various statistics such as the mean particle size, standard deviation of the PSD, etc. Plots of the PSD were produced using kernel density estimation procedures.[63] The data sets from different replications were combined to produce a single data set for each laboratory and each user.

All the remaining measurement methods produced size data in terms of frequency tables although their resolution differed. For example, differential mobility analysis (DMA) of NIST 8013 had bin sizes of around 1 nm, while the DLS frequency table for the same particle had bin sizes of various widths depending on the size of the center. These ranged from 1 nm to 20 nm. An example of this type of data for DMA is given in Table S3.

All methods produced replicated measurements. There are various methods of transforming this type of data into a PSD. The simplest method is to simply compute the proportions in each size category (i.e., bin), and then compute their averages and standard deviations. This approach can produce results that are not a true PSD in the sense that the averaged proportions do not have to add up to 1.

In this manuscript, the method used for calculating the PSD from a frequency table is based on a multinomial model.[64] The advantage of this method is that it always produces a PSD where the proportions add up to 1, and it comes with uncertainty that incorporates various sources, such as uncertainty due to repeatability, as well as uncertainty due to the resolution of the frequency table.

The multinomial statistical model states that the counts in the bins of the frequency table follow a multinomial distribution:

$$(n_{1j}, \dots, n_{kj}) \sim Multinomial(p_{1j}, \dots, p_{kj},\ N_j), j = 1, \dots, J \quad (5)$$

where J is the number of replicates, N_j is the total number of particles counted in replicate j, $(n_{1j}, \dots, n_{kj})$ are the particle counts in each bin in replicate j. The p_{ij} are the population proportions of particles in the i^{th} bin of replicate j. The objective is to estimate the p_{ij} and their uncertainties, and if necessary to combine them to obtain "average" values over the replicate samples. The average is not a simple arithmetic mean but is obtained using a hierarchical multinomial logit model[65] for the p_{ij} :

$$p_{ij} = \frac{\theta_{ij}}{\sum_{i=1}^{i=k} \theta_{ij}}, \theta_{ij} = e^{\beta_{ij}}, \beta_{ij} \sim N(\beta_{0i}, \sigma_i^2), i = 1, \dots, k, j = 1, \dots, J. \quad (6)$$

The notation $N(\beta_{0i}, \sigma_i^2)$ means a Gaussian distribution with mean β_{0i} and variance σ_i^2. In this statistical model the between-replicate uncertainty is represented by σ_i^2, and the "average" values of the proportions for category i are

$$\frac{e^{\beta_{0i}}}{\sum_{j=1}^{k} e^{\beta_{0j}}} \; i = 1, \dots, k. \quad (7)$$

To obtain point estimates and uncertainty of the relevant parameters we used a Bayesian analysis with non-informative prior distributions for the hyperparameters $\beta_{0i}, \sigma_i^2, i = 1, \dots, k$. We used a Gaussian distribution with a large variance (10^4) for the β_{0i}, and Inverse Gamma distribution with small shape parameters (10^{-5}) for σ_i^2. More information about Bayesian methods and prior distributions in metrological applications are provided in section 6.1 of reference[66]. The computations were done using Markov Chain Monte Carlo implemented in OpenBUGS.[67] Code is given in the Supplemental Methods.

Statistical Analysis: Derived Particle Number Concentration (PNC)

The derived NP number concentration formula for AuNPs is given as

$$PNC = \frac{C_{mass\ gold}}{\rho_{gold} \times \frac{\pi}{6} \times (size)^3} \quad (8)$$

where $C_{mass\ gold}$ is the Au mass concentration (μg/g), ρ_{gold} is the density of gold (19320 ± 1.4) kg/m^3 (uncertainty indicates standard uncertainty), *size* is the particle diameter (nm), and the units for PNC are particles/L. To determine $C_{mass\ gold}$ for each AuNP, the mean, the standard error of the mean, and the number of subsamples for individual total Au measurements were input into the NIST Consensus Builder (freely available at consensus.nist.gov) which applied the DerSimonian-Laird procedure described by Koepke et al.[68] to produce a consensus value for each AuNP with uncertainty bounds, and an estimate of the between-sample variability called dark uncertainty. The equation for deriving the PNC (8) is well defined under the condition that all particles are of the same diameter. In our case, there is a PSD for each particle type and measurement method. In what follows we treat PSD as a probability distribution of the random variable $size$. In this sense, PNC in (8) is a random variable with a probability distribution and an expected value

$$E[PNC] = E\left[\frac{C_{mass\ gold}}{\rho_{gold} \times \frac{\pi}{6} \times (\ size)^3}\right] = \sum \left(\frac{C_{mass\ gold}}{\rho_{gold} \times \frac{\pi}{6} \times (\ size)^3}\right) P(size)\,, \quad (9)$$

which we can use to represent PNC. Most often in the literature,[45, 51, 56, 57, 69] (8) has been approximated using the average (or expected) particle size $E(size)$ computed over the PSD as

$$PNC_{mean} = \frac{C_{mass\ gold}}{\rho_{gold} \times \frac{\pi}{6} \times (E(size))^3} \quad (10)$$

When the range of particle sizes is large, that is, when the variance of the PSD is large, or when the PSD is not symmetric, PNC_{mean} is not a very good approximation of *E*[*PNC*]. As described in subsequent sections, other features of the PSD may be better.

Because the function $\frac{C_{mass\ gold}}{\rho_{gold}\times\frac{\pi}{6}\times(E(size))^3}$ is convex in $size$, a mathematical property of expectation of convex functions, the Jensen's inequality,[70] guarantees that $PNC_{distribution} \geq PNC_{mean,}$ and so the estimate (10) is an underestimate of $E[PNC]$.

For measurement methods which produce individual particle sizes (spICP-MS), where the data is in the form of $s_1, \ldots, s_N$, $E[PNC]$ is computed simply as

$$E[PNC] = PNC_{distribution} = \frac{1}{N}\sum_{i=1}^{N}\frac{C_{mass\ gold}}{\rho_{gold}\times\frac{\pi}{6}\times(\ s_i)^3}\ . \qquad (11)$$

The specific value of $E[PNC]$ will depend on the sample size N, and in that sense, has an uncertainty associated with it. There is also uncertainty in $C_{mass\ gold}$ and ρ_{gold}, and so $E[PNC]$ has an uncertainty associated with it. This can be calculated using Monte Carlo propagation of uncertainty.[66] An example of the statistical models to include uncertainty due to $C_{mass\ gold}$ and ρ_{gold} for NIST RM 8013 is $C_{mass\ gold}\sim N(51.86, 0.32^2)$, $\rho_{gold}\sim N(19320, 1.4^2)$.

For size measurement methods that estimate PSD using a frequency table as in equation (6),

$$E[PNC] = PNC_{distribution} = \sum_{j=1}^{J}\sum_{i=1}^{k}\frac{C_{mass\ gold}\times p_{ij}}{\rho_{gold}\times\frac{\pi}{6}\times(\ s_i)^3}\ . \qquad (12)$$

When equation (7) is used, it becomes

$$E[PNC] = PNC_{distribution} = \sum_{i=1}^{k}\frac{C_{mass\ gold}\times\frac{e^{\beta_{0i}}}{\sum_{j=1}^{k}e^{\beta_{0j}}}}{\rho_{gold}\times\frac{\pi}{6}\times(\ s_i)^3}\ . \qquad (13)$$

As an example, the statistical model to estimate particle number concentration for NIST RM 8013 using DMA is:

$$C_{mass\ gold}\sim N(51.86, 0.32^2)$$

$$\rho_{gold}\sim N(19320, 1.4^2)$$

$$(n_1, \ldots, n_{13})_j \sim Multinomial\left(p_{1j}, \ldots, p_{13j},\ N_j\right), j = 1, \ldots, 20$$

$$p_{ij} = \frac{\theta_{ij}}{\sum_{i=1}^{i=13}\theta_{ij}},\ \theta_{ij} = e^{\beta_{ij}},\ \beta_{ij}\sim N(\beta_{0i}, \sigma_i^2)\ , i = 1, \ldots, 13, j = 1, \ldots, 20$$

$$PNC_{distribution} = \sum_{i=1}^{k=13}\frac{C_{mass\ gold}\times\frac{e^{\beta_{0i}}}{\sum_{j=1}^{k}e^{\beta_{0j}}}}{\rho_{gold}\times\frac{\pi}{6}\times(\ s_i)^3}\ .$$

Note that this statistical model assumes that all particles in bin i have diameters exactly equal to s_i. As the size of the bins ranges from 1.8 nm (bin center 47.8 nm) to 2.7 nm (bin center 73.7 nm), it is appropriate to account for this additional uncertainty in $PNC_{distribution}$ by letting size be a random variable with a triangular distribution as in $s_i\sim triangular(low, center, high)$ where low and $high$ are the two bin boundaries, and $center$ is the bin center.

The evaluation of this statistical model via Bayesian Markov Chain Monte Carlo produced our estimates of $PNC_{distribution}$ and of their uncertainty.

Statistical Analysis: Impact of Size Distribution Shape on Derived Particle Number Concentration (PNC)

Modeling of the impact of the shape of the distribution on the PNC was performed using a skew Normal distribution which is defined by the following equation:

$$f(x) = \frac{2}{s\sqrt{\pi}} e^{-\frac{(x-\theta)^2}{2s^2}} \int_{-\infty}^{a\left(\frac{x-\theta}{s}\right)} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\, dt \qquad (14)$$

with parameters θ (location), *s* (scale), and *a* (shape). When $a = 0$, the density becomes the Normal distribution. Skewness increases with the absolute value of *a*. Larger or smaller values of *s* correspond to greater or narrower widths of the distribution, respectively. Particle number concentrations were calculated for model distributions with different values of *a* and *s* using the approach described in the previous sections.

Results and Discussion

Gold mass concentration results

One key measurement used in the derivation of the PNC is the sample's total Au mass concentration. Thus, we first measured the total Au mass concentrations in all of the AuNP samples. For all samples, there was good agreement among all measurements of the total Au mass concentration performed using ICP-MS or ICP-optical emission spectroscopy (Figure S1). Surprisingly, one vial of the PVP AuNPs, which was opened at the beginning of the study, showed unexpected results with the PNC decreasing after every subsequent measurement (Figure S2). Therefore, we monitored this trend over a 2-year period, analyzing the sample periodically by both spICP-MS and total Au ICP-MS measurements. Interestingly, changes were not observed for unopened vials from the same manufacturer shipment or for vials of the bPEI AuNPs that had been open for a similar time period. While the cause of this change for the opened PVP AuNP sample was unclear, these results suggest that time-dependent changes in NP samples should be monitored to ensure that they do not impact total mass or PNC measurements.

The amount of Au measured in the process blanks was negligible. The mass fraction of Au measured in the supernatant of centrifuged samples, representing the "ionic" Au portion ranged from 30 µg/kg Au to 85 µg/kg which amounted to less than 0.2 % of the total Au mass fraction in the test materials. Analysis of the supernatant of centrifuged samples by spICP-MS showed that some AuNPs were present, indicating that the centrifuge procedure had not removed all of the AuNPs from suspension. As such, the "ionic" Au mass fraction in the test materials is no greater than 30 µg/kg Au to 85 µg/kg.

Size results

The values measured for the mean size in this study for DMA and SEM by Laboratory 2 for RM 8012 and 8013 were within 10 % of the values reported in the NIST Reports of Investigation (ROIs) (see Table S1),[54, 55] which falls within the expanded uncertainty of the

reference value (see Figures 1 and S3, or 2 for size distribution comparisons using boxplots or kernel density plots, respectively); kernel density plots show a smooth curve (i.e., without binning) estimating the probability density function of a continuous variability in this case the fraction of an AuNP over the size distribution. The DLS mean size values of RM 8012 and 8013 measured by Laboratory 3 and those measured in a recent study at NIST[71] of RM 8013 were also within 10 % of the ROI values.

A trend was found among all of the AuNPs analyzed which showed that the NTA and DLS analyses typically yielded larger size values than the techniques that only measured the core of the AuNP (e.g., spICP-MS and SEM), a result similar to other studies that compared NP size distributions among analytical techniques.[41-43] This finding could stem from a contribution of the surface coating and hydrated water ions to the size measurement for the NTA and DLS values or from these techniques being more strongly impacted by AuNP agglomerates. Overall, DLS provided the broadest size distributions among the techniques tested. This is likely due in part to DLS analysis being strongly impacted by the largest particles (signal intensity is proportional to diameter to the sixth power[72]), which has a substantial influence on polydisperse samples, and that the autocorrelation function measured by the instrument is deconvoluted using bins on a logarithmic scale and also assumptions of a monodisperse, normal distribution on the logarithmic scale. As a result of these limitations, output from DLS instruments is typically reported as the "z-average size" rather than a size distribution. However, the choice of which output to prioritize is case specific and depends on numerous factors such as the particle properties (e.g., geometry, dielectric constant, and size polydispersity) which impact the degree to which the sample can satisfy the Mie theory-based model and what is fit for purpose for the measurement.[72-75] In this study, we mainly used an intensity-based size distribution which follows the approach described in the NIST ROIs.[54, 55] The potential for agglomeration in DI water for each of the four AuNPs was evaluated using DLS (Figure S4). These analyses showed modest agglomeration across a 14 d period for the NIST RM 8012 and 8013 and PVP AuNP samples. In contrast, substantial agglomeration (initial peak decreased to near 0 %) was observed for the bPEI AuNP samples after 10 d or 4 d for the samples in plastic or glass, respectively (Figures S4 and S5).

One unexpected finding was the impact of the sample container (glass versus plastic) on the size measurements of the bPEI AuNP sample using hydrodynamic size-based techniques. From a visual inspection, it was clear that agglomeration rapidly occurred in the glass vials that were shipped from the manufacturer as indicated by a change in their color over time even when the samples were stored under refrigerated conditions (Figure S6). This finding was corroborated by DLS analyses which showed greater agglomeration for the samples shipped from the manufacturer in glass versus those shipped and stored in plastic (Figure S5). The different behaviors of the bPEI AuNPs in these different containers may be a result of different interactions of the AuNPs with the different container surfaces or the leaching of compounds from the containers which interacted with the AuNPs. The agglomeration of these samples, in addition to the potential for interactions of the positively-charged AuNPs with the different components of the sample introduction system, presented problems for several of the analytical techniques such as DMA, which was unable to measure the bPEI AuNP samples. While agglomeration of these samples also posed problems for NTA, adjusting the protocol (shortening the analysis period prior to redispersing the samples using

vortexing between runs) yielded results for the bPEI AuNP samples stored in plastic containers that were reproducible and exhibited a similar size distribution as the SEM and spICP-MS results (Figure 2). For the samples stored in glass, however, analysis using this revised protocol still yielded NTA size results that were substantially larger than those for spICP-MS and SEM with a tail that skewed toward larger sized particles (Figure S5). To evaluate to what extent the observed results could be impacted by vial-to-vial variability, three different vials of the bPEI AuNP samples in plastic were analyzed on the same day by spICP-MS; results indicated minimal vial-to-vial variability (Figure S7).

Modeling derived nanoparticle number concentration results

To better understand differences between the PNC_{mean} and $PNC_{distribution}$ values obtained from the five analytical techniques, statistical modeling was first performed to reveal the impact of various parameters on the PNC results. The results in this section reflect statistical modeling using skew Normal distributions, and not the size distributions measured in this study for any of the techniques. The theoretical impact of several parameters on the calculated NP number concentration was investigated: total Au mass concentration, mean of the NP size distribution (Θ), a tail in the distribution, and a change in the breadth of the distribution. Particle number concentrations were derived using $PNC_{distribution}$ or PNC_{mean} using equations (11) or (10), respectively.

The most straightforward parameter to evaluate was the impact of the gold concentration. A bias in this parameter was shown to have a linear impact on the NP number concentration (Figure 3A). For example, if the total gold concentration is underestimated by 5 % or 10 %, the NP number concentration will be similarly underestimated by 5 % or 10 %, respectively.

The impact of a bias in NP size (derived using either $PNC_{distribution}$ or PNC_{mean}) on the NP number concentration is more complex because it is asymmetric and based on size to the inverse third power. If the size is overestimated by a factor of 20 %, the PNC will be underestimated by ≈ 42 % (Figure 3B). Conversely, if the size is underestimated by a factor of 20 %, the PNC will be overestimated by ≈ 95 %. It is also important to point out that the magnitude of the bias in the NP number concentration exponentially decreases for smaller negative size biases. If the size is underestimated by 10 %, for example, the overestimation in NP number concentration is 37 %. Note that there were up to 10 % differences in the NP mean sizes measured by different techniques in this study compared to those measured in the NIST ROIs for RM 8012 and 8013, which would correspond to differences among PNC_{mean} values of – 24 % to + 37 %.

To evaluate the influence of the breadth of the NP distribution and skewed distributions with tails toward either smaller or larger NP sizes, fifteen model distributions were generated using skew Normal distributions (Figure 3C). When changing the breadth of the distribution (s), the impact on PNC_{mean} depended upon the skew (a). In the absence of a skew ($a = 0$), there was no impact on the PNC_{mean} (Figure 3D). When the skew resulted in a distribution with a tail toward larger particles ($a > 0$), increasing the breadth of the distribution decreased PNC_{mean}. This result stems from the tail toward larger particles increasing the mean size which would decrease the PNC

584 as shown in Figure 3B. The opposite trend was observed for distributions with a tail toward smaller
585 particles.

586 The trends for $PNC_{distribution}$ (Figure 3E) differed in some regards from those for PNC_{mean}.
587 When the value of *s* was greater than 0.05, the breadth of the distribution had a pronounced effect
588 on $PNC_{distribution}$ at some skewness values (*a* = -3, -2, or 0), but not when the distribution was
589 skewed toward larger particles (*a* = 2 or 3). For the distributions skewed toward larger particles,
590 there are counterbalancing trends: larger proportions of the distributions at larger NP sizes would
591 yield fewer particles, yet the broader distribution would also slightly increase the proportion of the
592 distribution at the smallest NP sizes, which would have a magnified impact on the derived PNC
593 (Figure 3B). For a distribution without a skew (*a* = 0) or a skew toward smaller particles (*a* = -3
594 or -2), broader distributions (greater *s* values) resulted in increased $PNC_{distribution}$ values (reflected
595 in the blue, green, and grey traces in Figure 3E). This result is similar to that of PNC_{mean} for
596 distributions with tails toward smaller particles, but the magnitude of the increase was
597 approximately a factor of five greater for $PNC_{distribution}$.

598 When directly comparing the modeled results for PNC_{mean} or $PNC_{distribution}$ (Figure 3F), one
599 trend is striking: regardless of the distribution, $PNC_{distribution}$ is always greater than PNC_{mean}. This
600 result is a consequence of the Jensen's inequality[70] since $PNC_{distribution}$ is a convex function of size,
601 and therefore, $PNC_{distribution} \geq PNC_{mean}$ (see Materials and methods section for additional details).
602 Overall, the magnitude of the difference between $PNC_{distribution}$ and PNC_{mean} increased with greater
603 breadth of the distribution (indicated by increasing *s* values) and for distributions with a greater
604 proportion of the distribution skewed toward smaller particles (i.e., the difference was smallest for
605 *a* = 2 or 3 and greatest for *a* = -2 or -3).

606 *Measured nanoparticle number concentration results*

607 Particle number concentrations were plotted for all the different particles using PNC_{mean},
608 $PNC_{distribution}$ or PNC_{direct} (Figure 4). In addition, pairwise comparisons (Figures 5, S8, S9, and S10)
609 were calculated among the PNC values for each AuNP. This allowed for a direct comparison
610 among techniques in terms of the degree of difference between their results, and also among the
611 results for each technique depending upon if the mean or PSD was used to derive the PNC value,
612 or if PNC was directly measured by the instrument. In agreement with the modeling, $PNC_{distribution}$
613 (indicated by an orange marker in Figure 4) was greater than PNC_{mean} (indicated by a purple
614 marker) for all conditions tested. The $PNC_{distribution}$ and PNC_{mean} values were closest for the
615 distributions without a tail toward smaller particles and for narrower distributions (e.g., the SEM
616 results for RM 8012 or 8013) (Figures 5, 6, S9, S10, and S11), a result also in agreement with the
617 modeling. The biggest discrepancies between the $PNC_{distribution}$ and PNC_{mean} values were typically
618 observed for the techniques that yielded the broadest distributions such as DLS. Importantly, these
619 results reveal that this seemingly unimportant choice, namely whether to calculate $PNC_{distribution}$ or
620 PNC_{mean} values, can have a substantive (potentially > 50 %) impact on the derived PNC.

621 The influence of estimating the PNC using a range of central tendency indicators (mean,
622 median, mode, 10 % trimmed mean, 10 % winsorized mean, and $M_{estimator}$) or $PNC_{distribution}$ were
623 compared for the PVP AuNP sample (Figure 7; these central tendency indicators are defined in

Table S2). The PVP AuNP sample was chosen since it was more polydisperse than the NIST RMs yet was able to be analyzed by all techniques, unlike for the bPEI AuNPs. While the use of alternate central tendency indicators to derive the PNC typically yielded PNC results that were less than $PNC_{distribution}$, there were some scenarios, for example using the mode as the central tendency indicator for DMA data, where PNCs were greater than $PNC_{distribution}$ (Figure 7A). For all techniques except for DLS, the difference between the PNC derived using the different central tendency indicators and $PNC_{distribution}$ was less than 20 % (Figure 7B). The greater difference for DLS between PNC values derived using central tendency indicators or the full size distribution can be explained by the substantially broader size distribution for DLS compared to those for the other techniques (Figure 2C); this resulted in $PNC_{distribution}$ values for DLS that were more strongly impacted by the tail of the distribution toward smaller particles. Overall, the central tendency indicator that yielded results closest to $PNC_{distribution}$ was the mode (Figure 7B). However, it was unclear to what extent this result would be generalizable to other samples since it cannot be explained by a mathematical formula.

For the techniques that provided PNC_{direct} values (spICP-MS and NTA), it is informative to compare these values to $PNC_{distribution}$ values obtained from the same technique (Figures 4 and 5). For spICP-MS measurements, PNC_{direct} values for all of the samples except for the bPEI sample stored in glass are 3 % to 31 % lower than $PNC_{distribution}$ measured by this technique (Figures 5, S8, S9, and S10). This result may stem from NP losses within the sample introduction system for the PNC_{direct} measurements or the impact of the instrument calibration procedure and in particular the calculation of the transport efficiency.[51] While it is possible $PNC_{distribution}$ may be overestimated if the density for the NPs is lower than that of the bulk metal,[76] this is unlikely to bias the measurements reported here since the calibration and PNC_{direct} measurements used the same density value (i.e., that of bulk Au). In contrast to the spICP-MS results, $PNC_{distribution}$ values for NTA were less than those directly measured using this technique for RM 8012 and the PVP and bPEI AuNPs (Figure 4A, 4C and 4D). This result likely stems from the size distribution measured using NTA being shifted to larger particles compared to those for most other techniques, which would result in a relatively lower $PNC_{distribution}$.

One valuable approach for comparing among NP size measurement techniques is evaluation against an established reference technique such as electron microscopy. In this study, $PNC_{distribution}$, PNC_{mean}, and PNC_{direct} results for all techniques were compared to $PNC_{distribution}$ values measured in Laboratory 1 using SEM (Figure 6); the NIST ROI size distribution values were used for the 8012 and 8013 samples, while those for the PVP and bPEI AuNP samples were independently measured. Similar to the results from the NP size distribution measurements (Figures 1 and 2), the $PNC_{distribution}$ values from the spICP-MS from both laboratories and SEM results from Laboratory 2 were generally the closest to the SEM value (indicated by the dotted grey line in Figure 4 which shows $PNC_{distribution}$ for the Laboratory 1 SEM results). $PNC_{distribution}$ and PNC_{mean} results from DMA, NTA, and DLS measurements often differed by greater than 30 % compared to the $PNC_{distribution}$ values using SEM (Figure 6). PNC_{direct} using spICP-MS were 5 % to 26 % less than $PNC_{distribution}$ using SEM for the samples tested. There was not a consistent trend between PNC_{direct} by NTA and $PNC_{distribution}$ using SEM with PNC_{direct} using NTA being

substantially (34 % to 36 %) less for the two RM AuNPs yet 5 % or 40 % greater for the PVP and bPEI AuNPs, respectively.

One interesting result from the pairwise comparison is the differing results for the direct measurements of the bPEI sample in the two different containers using spICP-MS. For laboratory 1, PNC_{direct} and $PNC_{distribution}$ were within 30 % using spICP-MS for the bPEI AuNPs in plastic containers, yet ranged from 46 to 57 % for the bPEI AuNPs in glass containers. In all cases, results for PNC_{direct} were lower than $PNC_{distribution}$. It is possible that the overall lower recovery for PNC_{direct} as compared to $PNC_{distribution}$ for measurements performed by spICP-MS may indicate loss of material to sample containers across the high dilution needed to properly perform an spICP-MS analysis. That the total Au mass fractions measured for the stock suspensions were close to the expected values provided by the supplier indicates that any AuNP loss occurred in subsequent dilution steps. Additional particle loss can occur within the sample introduction system of the ICP-MS (i.e. transport tubing, nebulizer and spray chamber). However, the spICP-MS PNC_{direct} measurements for the bPEI AuNPs in glass containers showed the lowest recovery. This may serve as further evidence of the impact of the glass storage container on the bPEI AuNPs and supports the observation that storage of bPEI AuNPs in glass caused agglomeration with resulting lower number concentration. Importantly, these samples yielded similar size distributions using spICP-MS and SEM, although the spICP-MS distribution for the samples in glass containers was broader (Figure S5). Clearly, the ability of spICP-MS to measure both PNC_{direct} and $PNC_{distribution}$ assists in understanding the unique behavior of each nanoparticle system.

Impact on in vitro NP dosimetry

The impact of using the PSD or the mean diameter and of using different analytical techniques on predicted cell dosimetry was evaluated using the Distorted Grid version of the *In vitro* Sedimentation, Diffusion, and Dosimetry model (DG-ISDD).[62, 77] While it is possible in this model to adjust the adsorption properties ("stickiness") of the lower boundary condition reflecting the potential for different NPs to be associated with the cell surface to variable extents,[77] the modeling performed in this paper assumed a perfectly adsorptive boundary condition. When evaluating the impact of different input parameters to ISDD models, the influence from uncertainty of measuring the NP size distribution has been generally treated as being modest relative to other sources of uncertainty.[62, 77, 78] However, the impact of different NP size measurement techniques on the modeled *in vitro* cellular concentration has not yet been evaluated. The modeled values for different size measurement techniques, which were calculated using either the mean or PSD for each technique, were relatively close ($\approx > 10$ %) for some samples (e.g., RM 8012), yet differed substantially (a factor of 3) for the bPEI sample in glass (Figure 8).

To better understand the DG-ISDD results, the modeled *in vitro* concentrations across a range of AuNP sizes were modeled (Figure S11). The largest AuNPs (80 nm and 90 nm) showed nearly complete association with the cells after 24 h as a result of sedimentation. It is interesting to note that the amount of deposited AuNPs was nearly identical for the 20 nm and 30 nm AuNPs, while there was a ≈ 20 % greater deposition fraction for the 40 nm AuNPs. Therefore, for the ≈ 30 nm AuNPs, a tail toward smaller particles for a symmetrical distribution would not have as much of an impact on the cellular concentration as would the tail toward larger particles. This was likely

a result of a decreasing impact of diffusion on the *in vitro* concentration counterbalanced by an increasing effect from sedimentation in this size range. For AuNPs with a size of ≈ 60 nm, there would be a similar magnitude of an impact on the deposited fraction for tails toward smaller or larger sized AuNPs with tails toward smaller AuNPs yielding less deposition while tails toward larger AuNPs would have greater deposition rates.

To compare the influence of using the mean or the PSD when calculating the deposited cellular concentration, comparisons were made for each technique (Figure S12A). Overall, the magnitude of the difference between using the mean or PSD was frequently less than 10 %, although larger differences were observed for some samples for DLS and NTA. This result is similar to that from a previous study which showed that polydisperse samples could have substantially different cellular concentrations depending on whether the mean or PSD was used in the modeling.[77] When comparing the cellular concentration for each technique to the results for the PSD for SEM (Figure S12B), there were substantially greater differences which in some cases were up to 65 %. Importantly, the greatest difference among the techniques was observed for the bPEI AuNP sample in glass, the sample that showed the largest amount of agglomeration. Since NPs often agglomerate extensively in cellular media,[79-81] this suggests that methods to improve the precision of size distribution measurements of agglomerated NPs would help decrease the uncertainty in the modeled cellular concentration, because it is challenging to accurately measure samples with broad PSDs including large agglomerates.

In addition to performing modeling using the intensity-based DLS distributions which are utilized throughout the manuscript, modeling was also performed using the volume-based DLS distribution as suggested by Deloid et al.[62] Dynamic light scattering distributions can typically be derived using intensity-, volume-, or number-based distributions, and which distribution to use depends upon case specific criteria.[82] The volume-based distributions yielded predicted deposited percentages that were 4 %, 29 %, and 24 % lower than the intensity-based distributions for the RM 8012, 8013, and PVP AuNP samples, respectively. Calculations were not performed for the bPEI samples since the volume-based size distribution yielded unrealistic results. Overall, full size distributions using DLS often differed substantially from those measured using the other techniques, and a recent framework for characterizing nanoparticles for medical applications discouraged the use of DLS for measuring size distributions.[83] Using DLS for measuring the full size distribution as the input for NP dosimetry modeling may lead to results that substantially differ from those calculated using other high resolution techniques.

Conclusions

This multi-method analytical and modeling evaluation of PNC has yielded valuable findings regarding recommendations for the usage of PNCs in future research and decision making. Employing $PNC_{distribution}$ has several advantages over using PNC_{mean} for PNC measurements, because PNC_{mean} is not a very good approximation of $PNC_{distribution}$ when the PSD is broad especially with tails toward smaller particles, not symmetric, or bimodal.[58] However, the uncertainty for the percentage of the distribution in the tails would typically be larger than that for a central tendency indicator such as the mean, and the PSD is substantially more challenging to accurately calculate for some techniques such as DLS. Underestimation of the NP size or of the

percentage of a PSD in a tail toward smaller NP sizes would have a magnified influence on $PNC_{distribution}$ compared to PNC_{mean} values. Therefore, improving the accuracy of measurements of the PSD and the comparability of results among analytical methods are key topics for ongoing research. This would also support the increased use of PNC values in nanoecotoxicology research. When comparing $PNC_{distribution}$ results, the techniques which only measured the core of the NP, namely spICP-MS and SEM, were in closer agreement to each other than the results from the other techniques. DLS typically yielded results that differed most substantially from the other techniques. Given the widespread adoption of some techniques such as DLS, it is critical to recognize its limitations with regards to deriving a PNC from the measured DLS PSD. Analysis of the PVP and bPEI AuNP samples yielded more variable results than those from the RM samples. For the positively-charged bPEI AuNPs, many techniques were unable to yield reliable results. Therefore, additional research is recommended to improve the characterization of NPs with different surface coatings. Overall, the differences observed among techniques suggest it would be helpful to improve the agreement of these methods prior to usage of PNC values for regulatory testing using OECD methods especially since some samples will contain agglomerated NPs which is expected to further increase the difference in the values obtained among techniques. Lastly, this study revealed that the analytical method chosen to measure the PSD can also have a substantial impact on the modeled cellular concentration. Thus, the choice of which analytical technique was used to measure the size distribution could in some cases yield substantially different modeled concentrations that reaches the cells, thus potentially altering interpretations of the results.

Acknowledgements

We acknowledge funding from the NanoScreen Materials Challenge co-funded by the Competence Centre for Materials Science and Technology (CCMX). We thank NIST colleagues Kavuri P. Purushotham and András E. Vladár for their contribution to the SEM analysis.

Conflicts of interest

There are no conflicts of interest to declare.

Author contributions

A.R.M.B., M.J., M. E., G.C.C., A.L.N., Q.C., J.K., B.M., K.M., and M.R. performed the experimental measurements. E.J.P. performed the DG-ISDD modeling. B.M. performed the statistical analyses. E.J.P. wrote the manuscript with contributions from the other authors.

References

1. J. R. Potts, D. R. Dreyer, C. W. Bielawski and R. S. Ruoff, Graphene-based polymer nanocomposites, *Polymer*, 2011, **52**, 5-25.
2. S. Pavlidou and C. D. Papaspyrides, A review on polymer-layered silicate nanocomposites, *Prog. Polymer Sci.*, 2008, **33**, 1119-1198.
3. H. Zou, S. S. Wu and J. Shen, Polymer/silica nanocomposites: Preparation, characterization, properties, and applications, *Chem. Rev.*, 2008, **108**, 3893-3957.
4. T. Nguyen, E. J. Petersen, B. Pellegrin, J. M. Gorham, T. Lam, M. Zhao and L. Sung, Impact of UV irradiation on multiwall carbon nanotubes in nanocomposites: Formation of entangled surface layer and mechanisms of release resistance, *Carbon*, 2017, **116**, 191-200.
5. E. J. Petersen, R. A. Pinto, X. Y. Shi and Q. G. Huang, Impact of size and sorption on degradation of trichloroethylene and polychlorinated biphenyls by nano-scale zerovalent iron, *J. Hazard. Mater.*, 2012, **243**, 73-79.
6. ISO (International Organization for Standardization), TS 80004-1: nanotechnologies — vocabulary — Part 1: Core terms. 2010.
7. ASTM (American Society for Testing Materials) International, E2456-06: standard terminology relating to nanotechnology. 2006.
8. 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**, 2108-2116.3, **7**, 21-29.
9. B. Nowack, R. M. David, H. Fissan, H. Morris, J. A. Shatkin, M. Stintz, R. Zepp and D. Brouwer, Potential release scenarios for carbon nanotubes used in composites, *Environ. Intl.*, 2013, **59**, 1-11.
10. D. M. Mitrano, S. Motellier, S. Clavaguera and B. Nowack, Review of nanomaterial aging and transformations through the life cycle of nano-enhanced products, *Environ. Intl.*, 2015, **77**, 132-147.
11. S. J. Froggett, S. F. Clancy, D. R. Boverhof and R. A. Canady, A review and perspective of existing research on the release of nanomaterials from solid nanocomposites, *Part. Fibre Toxicol.*, 2014, **11**.
12. D. Singh, G. A. Sotiriou, F. Zhang, J. Mead, D. Bello, W. Wohlleben and P. Demokritou, End-of-life thermal decomposition of nano-enabled polymers: effect of nanofiller loading and polymer matrix on by-products, *Environ. Sci.: Nano*, 2016, **3**, 1293-1305.
13. R. Bjorkland, D. A. Tobias and E. J. Petersen, Increasing evidence indicates low bioaccumulation of carbon nanotubes, *Environ. Sci.: Nano*, 2017, **4**, 747-766.
14. M. Mortimer, E. J. Petersen, B. A. Buchholz, E. Orias and P. A. Holden, Bioaccumulation of Multiwall Carbon Nanotubes in Tetrahymena thermophila by Direct Feeding or Trophic Transfer, *Environ. Sci. Technol.*, 2016, **50**, 8876-8885.
15. M. E. Johnson, S. K. Hanna, A. R. M. Bustos, C. M. Sims, L. C. C. Elliott, A. Lingayat, A. C. Johnston, B. Nikoobakht, J. T. Elliott, R. D. Holbrook, K. C. K. Scoto, K. E. Murphy, E. J. Petersen, L. L. Yu and B. C. Nelson, Separation, Sizing, and Quantitation of Engineered Nanoparticles in an Organism Model Using Inductively Coupled Plasma Mass Spectrometry and Image Analysis, *ACS Nano*, 2017, **11**, 526-540.
16. 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.
17. A. E. Nel, E. Nasser, H. Godwin, D. Avery, T. Bahadori, L. Bergeson, E. Beryt, J. C. Bonner, D. Boverhof, J. Carter, V. Castranova, J. R. DeShazo, S. M. Hussain, A. B. Kane, F. Klaessig, E.

Kuempel, M. Lafranconi, R. Landsiedel, T. Malloy, M. B. Miller, J. Morris, K. Moss, G. Oberdorster, K. Pinkerton, R. C. Pleus, J. A. Shatkin, R. Thomas, T. Tolaymat, A. Wang and J. Wong, A Multi-Stakeholder Perspective on the Use of Alternative Test Strategies for Nanomaterial Safety Assessment, *ACS Nano*, 2013, **7**, 6422-6433.

18. J. A. Shatkin, K. J. Ong, C. Beaudrie, A. J. Clippinger, C. O. Hendren, L. T. Haber, M. Hill, P. Holden, A. J. Kennedy, B. Kim, M. MacDonell, C. M. Powers, M. Sharma, L. Sheremeta, V. Stone, Y. Sultan, A. Turley and R. H. White, Advancing Risk Analysis for Nanoscale Materials: Report from an International Workshop on the Role of Alternative Testing Strategies for Advancement, *Risk Analysis*, 2016, **36**, 1520-1537.
19. A. J. Clippinger, A. Ahluwalia, D. Allen, J. C. Bonner, W. Casey, V. Castranova, R. M. David, S. Halappanavar, J. A. Hotchkiss, A. M. Jarabek, M. Maier, W. Polk, B. Rothen-Rutishauser, C. M. Sayes, P. Sayre, M. Sharma and V. Stone, Expert consensus on an in vitro approach to assess pulmonary fibrogenic potential of aerosolized nanomaterials, *Arch. Toxicol.*, 2016, **90**, 1769-1783.
20. M. Sharma, J. Nikota, S. Halappanavar, V. Castranova, B. Rothen-Rutishauser and A. J. Clippinger, Predicting pulmonary fibrosis in humans after exposure to multi-walled carbon nanotubes (MWCNTs), *Arch. Toxicol.*, 2016, **90**, 1605-1622.
21. S. K. Hanna, G. A. Cooksey, S. Dong, B. C. Nelson, L. Mao, J. T. Elliott and E. J. Petersen, Feasibility of using a standardized Caenorhabditis elegans toxicity test to assess nanomaterial toxicity, *Environ. Sci.: Nano*, 2016, **3**, 1080-1089.
22. E. J. Petersen, D. X. Flores-Cervantes, T. D. Bucheli, L. C. C. Elliott, J. A. Fagan, A. Gogos, S. Hanna, R. Kägi, E. Mansfield, A. R. M. Bustos, D. L. Plata, V. Reipa, P. Westerhoff and M. R. Winchester, Quantification of Carbon Nanotubes in Environmental Matrices: Current Capabilities, Case Studies, and Future Prospects, *Environ. Sci. Technol.*, 2016, **50**, 4587-4605.
23. H. Selck, R. D. Handy, T. F. Fernandes, S. J. Klaine and E. J. Petersen, Nanomaterials in the aquatic environment: A European Union–United States perspective on the status of ecotoxicity testing, research priorities, and challenges ahead, *Environ. Toxicol. Chem.*, 2016, **35**, 1055-1067.
24. M. D. Montano, G. V. Lowry, F. von der Kammer, J. Blue and J. F. Ranville, Current status and future direction for examining engineered nanoparticles in natural systems, *Environ. Chem.*, 2014, **11**, 351-366.
25. F. Laborda, E. Bolea, G. Cepria, M. T. Gomez, M. S. Jimenez, J. Perez-Arantegui and J. R. Castillo, Detection, characterization and quantification of inorganic engineered nanomaterials: A review of techniques and methodological approaches for the analysis of complex samples, *Anal. Chem. Acta*, 2016, **904**, 10-32.
26. F. Laborda, J. Jimenez-Lamana, E. Bolea and J. R. Castillo, Critical considerations for the determination of nanoparticle number concentrations, size and number size distributions by single particle ICP-MS, *J. Anal. At. Spectrom.*, 2013, **28**, 1220-1232.
27. D. G. Goodwin, A. S. Adeleye, L. Sung, K. T. Ho, R. M. Burgess and E. J. Petersen, Detection and Quantification of Graphene-Family Nanomaterials in the Environment, *Environ. Sci. Technol.*, 2018, **52**, 4491-4513.
28. F. A. Monikh, L. Chupani, E. Zuskova, R. Peters, M. Vancova, M. G. Vijver, P. Porcal and W. Peijnenburg, Method for Extraction and Quantification of Metal-Based Nanoparticles in Biological Media: Number-Based Biodistribution and Bioconcentration, *Environ. Sci. Technol.*, 2019, **53**, 946-953.
29. G. Oberdörster, E. Oberdörster and J. Oberdörster, Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles, *Environ. Health Perspect.*, 2005, **113**, 823-839.

30. L. Lagier, F. Mouchet, C. Laplanche, A. Mottier, S. Cadarsi, L. Evariste, C. Sarrieu, P. Lonchambon, E. Pinelli, E. Flahaut and L. Gauthier, Surface area of carbon-based nanoparticles prevails on dispersion for growth inhibition in amphibians, *Carbon*, 2017, **119**, 72-81.
31. A. Mottier, F. Mouchet, C. Laplanche, S. Cadarsi, L. Lagier, J. C. Arnault, H. A. Girard, V. Leon, E. Vazquez, C. Sarrieu, E. Pinelli, L. Gauthier and E. Flahaut, Surface Area of Carbon Nanoparticles: A Dose Metric for a More Realistic Ecotoxicological Assessment, *Nano Lett.*, 2016, **16**, 3514-3518.
32. A. J. Kennedy, M. S. Hull, S. Diamond, M. Chappell, A. J. Bednar, J. G. Laird, N. L. Melby and J. A. Steeyens, Gaining a Critical Mass: A Dose Metric Conversion Case Study Using Silver Nanoparticles, *Environ. Sci. Technol.*, 2015, **49**, 12490-12499.
33. K. Van Hoecke, K. A. C. De Schamphelaere, P. Van der Meeren, S. Lucas and C. R. Janssen, Ecotoxicity of silica nanoparticles to the green alga *Pseudokirchneriella subcapitata*: Importance of surface area, *Environ. Toxicol. Chem.*, 2008, **27**, 1948-1957.
34. Y. Deng, E. J. Petersen, K. E. Challis, S. A. Rabb, R. D. Holbrook, J. F. Ranville, B. C. Nelson and B. Xing, Multiple Method Analysis of TiO2 Nanoparticle Uptake in Rice (Oryza sativa L.) Plants, *Environ. Sci. Technol.*, 2017, **51**, 10615-10623.
35. E. J. Petersen, M. Mortimer, R. M. Burgess, R. Handy, S. Hanna, K. T. Ho, M. Johnson, S. Loureiro, H. Selck, J. J. Scott-Fordsmand, D. Spurgeon, J. Unrine, N. W. van den Brink, Y. Wang, J. White and P. Holden, Strategies for robust and accurate experimental approaches to quantify nanomaterial bioaccumulation across a broad range of organisms, *Environ. Sci.: Nano*, 2019, **6**, 1619-1656.
36. F. Abdolahpur Monikh, L. Chupani, M. G. Vijver, M. Vancová and W. J. G. M. Peijnenburg, Analytical approaches for characterizing and quantifying engineered nanoparticles in biological matrices from an (eco)toxicological perspective: old challenges, new methods and techniques, *Sci. Tot. Environ.*, 2019, **660**, 1283-1293.
37. F. Abdolahpur Monikh, L. Chupani, E. Zusková, R. Peters, M. Vancová, M. G. Vijver, P. Porcal and W. J. G. M. Peijnenburg, Method for Extraction and Quantification of Metal-Based Nanoparticles in Biological Media: Number-Based Biodistribution and Bioconcentration, *Environ. Sci. Technol.*, 2019, **53**, 946-953.
38. S. C. Brown, V. Boyko, G. Meyers, M. Voetz and W. Wohlleben, Toward Advancing Nano-Object Count Metrology: A Best Practice Framework, *Environ. Health Perspect.*, 2013, **121**, 1282-1291.
39. European Commission. 2011/696/EU: Commission Recommendation of 18 October 2011 on the definition of nanomaterial. Off. J. Eur. Communities: Legis. 2011, 275, 38−40.
40. E. J. Petersen, S. A. Diamond, A. J. Kennedy, G. G. Goss, K. Ho, J. Lead, S. K. Hanna, N. B. Hartmann, K. Hund-Rinke, B. Mader, N. Manier, P. Pandard, E. R. Salinas and P. Sayre, Adapting OECD Aquatic Toxicity Tests for Use with Manufactured Nanomaterials: Key Issues and Consensus Recommendations, *Environ. Sci. Technol.*, 2015, **49**, 9532-9547.
41. H. E. Pace, N. J. Rogers, C. Jarolimek, V. A. Coleman, E. P. Gray, C. P. Higgins and J. F. Ranville, Single Particle Inductively Coupled Plasma-Mass Spectrometry: A Performance Evaluation and Method Comparison in the Determination of Nanoparticle Size, *Environ. Sci. Technol.*, 2012, **46**, 12272-12280.
42. R. F. Domingos, M. A. Baalousha, Y. Ju-Nam, M. M. Reid, N. Tufenkji, J. R. Lead, G. G. Leppard and K. J. Wilkinson, Characterizing Manufactured Nanoparticles in the Environment: Multimethod Determination of Particle Sizes, *Environ. Sci. Technol.*, 2009, **43**, 7277-7284.
43. J. S. Taurozzi, V. A. Hackley and M. R. Wiesner, A standardised approach for the dispersion of titanium dioxide nanoparticles in biological media, *Nanotoxicology*, 2013, **7**, 389-401.

44. A. Dudkiewicz, A. B. Boxall, Q. Chaudhry, K. Molhave, K. Tiede, P. Hofmann and T. P. Linsinger, Uncertainties of size measurements in electron microscopy characterization of nanomaterials in foods, *Food Chem*, 2015, **176**, 472-479.
45. A. R. Montoro Bustos, E. J. Petersen, A. Possolo and M. R. Winchester, Post hoc Interlaboratory Comparison of Single Particle ICP-MS Size Measurements of NIST Gold Nanoparticle Reference Materials, *Anal. Chem.*, 2015, **87**, 8809-8817.
46. B. T. Mader, M. E. Ellefson and S. T. Wolf, Measurements of nanomaterials in environmentally relevant water matrices using liquid nebulization/differential mobility analysis, *Environ. Toxicol. Chem.*, 2015, **34**, 833-842.
47. C. M. Maguire, K. Sillence, M. Roesslein, C. Hannell, G. Suarez, J. J. Sauvain, S. Capracotta, S. Contal, S. Cambier, N. El Yamani, M. Dusinska, A. Dybowska, A. Vennemann, L. Cooke, A. Haase, A. Luch, M. Wiemann, A. Gutleb, R. Korenstein, M. Riediker, P. Wick, P. Hole and A. Prina-Mello, Benchmark of Nanoparticle Tracking Analysis on Measuring Nanoparticle Sizing and Concentration, *J.of Micro and Nano-Manufact.*, 2017, **5**.
48. T. P. Linsinger, R. Peters and S. Weigel, International interlaboratory study for sizing and quantification of Ag nanoparticles in food simulants by single-particle ICPMS, *Anal. Bioanal. Chem.*, 2014, **406**, 3835-3843.
49. S. Weigel, R. Peters, K. Loeschner, R. Grombe and T. P. J. Linsinger, Results of an interlaboratory method performance study for the size determination and quantification of silver nanoparticles in chicken meat by single-particle inductively coupled plasma mass spectrometry (sp-ICP-MS), *Anal. Bioanal. Chem.*, 2017, **409**, 4839-4848.
50. R. Sekine, K. Khurana, K. Vasilev, E. Lombi and E. Donner, Quantifying the adsorption of ionic silver and functionalized nanoparticles during ecotoxicity testing: Test container effects and recommendations, *Nanotoxicology*, 2015, **9**, 1005-1012.
51. J. Y. Liu, K. E. Murphy, M. R. Winchester and V. A. Hackley, Overcoming challenges in single particle inductively coupled plasma mass spectrometry measurement of silver nanoparticles, *Anal. Bioanal. Chem.*, 2017, **409**, 6027-6039.
52. A. K. Pal, D. Bello, J. Cohen and P. Demokritou, Implications of in vitro dosimetry on toxicological ranking of low aspect ratio engineered nanomaterials, *Nanotoxicology*, 2015, **9**, 871-885.
53. M. K. Ha, Y. J. Shim and T. H. Yoon, Effects of agglomeration on *in vitro* dosimetry and cellular association of silver nanoparticles, *Environ. Sci.: Nano*, 2018, **5**, 446-455.
54. NIST, Reference Material® 8012 Gold Nanoparticles, Nominal 30 nm Diameter. 2015.
55. NIST, Reference Material® 8013 Gold Nanoparticles, Nominal 60 nm Diameter. 2015.
56. H. E. Pace, N. J. Rogers, C. Jarolimek, V. A. Coleman, C. P. Higgins and J. F. Ranville, Determining Transport Efficiency for the Purpose of Counting and Sizing Nanoparticles via Single Particle Inductively Coupled Plasma Mass Spectrometry, *Anal. Chem.*, 2011, **83**, 9361-9369.
57. K. E. Murphy, J. Lui, A. R. Montoro Bustos, M. E. Johnson and M. R. Winchester, NIST Special Publication 1200-21: Characterization of nanoparticle suspensions using single particle inductively coupled plasma mass spectrometry. Version 1.0. 2015, DOI: doi:10.6028/NIST.SP.1200-21
58. A. R. Montoro Bustos, K. P. Purushotham, A. Possolo, N. Farkas, A. E. Vladár, K. E. Murphy and M. R. Winchester, Validation of Single Particle ICP-MS for Routine Measurements of Nanoparticle Size and Number Size Distribution, *Anal. Chem.*, 2018, **90**, 14376-14386.
59. H. Saveyn, B. De Baets, O. Thas, P. Hole, J. Smith and P. Van der Meeren, Accurate particle size distribution determination by nanoparticle tracking analysis based on 2-D Brownian dynamics simulation, *J. Coll. Interf. Sci.*, 2010, **352**, 593-600.

60. D. H. Tsai, R. A. Zangmeister, L. F. Pease, 3rd, M. J. Tarlov and M. R. Zachariah, Gas-phase ion-mobility characterization of SAM-functionalized Au nanoparticles, *Langmuir*, 2008, **24**, 8483-8490.
61. A. E. Vladar and B. Ming, NIST-NCL joint assay protocol PCC-15: Measuring the size of colloidal gold nano-particles using high-resolution scanning electron microscopy. 2011, p. 20.
62. G. M. DeLoid, J. M. Cohen, G. Pyrgiotakis and P. Demokritou, Preparation, characterization, and in vitro dosimetry of dispersed, engineered nanomaterials, *Nat. Prot.*, 2017, **12**, 355-371.
63. Y.-C. Chen, A tutorial on kernel density estimation and recent advances, *Biostatistics & Epidemiology*, 2017, **1**, 161-187.
64. M. Evans, N. Hastings and B. Peacock, in *Statistical distributions 3rd ed.*, 2000, pp. 134-136.
65. Z. Chen and L. Kuo, A Note on the Estimation of the Multinomial Logit Model With Random Effects, *Amer. Statistic.*, 2001, **55**, 89-95.
66. A. Possolo and B. Toman, Tutorial for metrologists on the probabilistic and statistical apparatus underlying the GUM and related documents. National Institute of Standards and Technology, Gaithersburg, MD, November 2011. doi: 10.13140/RG.2.1.2256.8482. URL www.itl.nist.gov/div898/possolo/TutorialWEBServer/TutorialMetrologists2011Nov09.xht.).
67. L. David, S. David, T. Andrew and B. Nicky, The BUGS project: Evolution, critique and future directions, *Stat. Medic.*, 2009, **28**, 3049-3067.
68. A. Koepke, T. Lafarge, A. Possolo and B. Toman, Consensus building for interlaboratory studies, key comparisons, and meta-analysis, *Metrologia*, 2017, **54**, 534-562.
69. A. R. Montoro Bustos, J. M. Pettibone and K. E. Murphy, in *Nanoparticle Design and Characterization for Catalytic Applications in Sustainable Chemistry*, The Royal Society of Chemistry, 2019, DOI: 10.1039/9781788016292-00037, pp. 37-83.
70. G. Hardy, J. Littlewood and G. Polya, in *Inequalities, 2nd ed.*, Cambridge University Press, Cambridge, England, 1988, pp. 83-84.
71. S. K. Hanna, A. R. Montoro Bustos, A. W. Peterson, V. Reipa, L. D. Scanlan, S. Hosbas Coskun, T. J. Cho, M. E. Johnson, V. A. Hackley, B. C. Nelson, M. R. Winchester, J. T. Elliott and E. J. Petersen, Agglomeration of Escherichia coli with Positively Charged Nanoparticles Can Lead to Artifacts in a Standard Caenorhabditis elegans Toxicity Assay, *Environ. Sci. Technol.*, 2018, **52**, 5968-5978.
72. J. Stetefeld, S. A. McKenna and T. R. Patel, Dynamic light scattering: a practical guide and applications in biomedical sciences, *Biophys. Rev.*, 2016, **8**, 409-427.
73. S. P. Pan, H. F. Weng, C. M. Lin and T. S. Liu, Uncertainty Analysis on Precision Measurement for Polystyrene Nanospheres Using Dynamic Light Scattering, *Jap. J. Appl. Phys.*, 2010, **49**.
74. D. Biganzoli and F. Ferri, Statistical analysis of dynamic light scattering data: revisiting and beyond the Schatzel formulas, *Optics Express*, 2018, **26**, 29375-29392.
75. S. Aleandri, A. Vaccaro, R. Id, Armenta, A. Charles Völker and M. Kuentz, *Dynamic Light Scattering of Biopharmaceutics—Can Analytical Performance Be Enhanced by Laser Power?*, 2018.
76. S. Tadjiki, M. D. Montano, S. Assemi, A. Barber, J. Ranville and R. Beckett, Measurement of the Density of Engineered Silver Nanoparticles Using Centrifugal FFF-TEM and Single Particle ICP-MS, *Anal. Chem.*, 2017, **89**, 6057-6065.
77. G. M. DeLoid, J. M. Cohen, G. Pyrgiotakis, S. V. Pirela, A. Pal, J. Liu, J. Srebric and P. Demokritou, Advanced computational modeling for in vitro nanomaterial dosimetry, *Part. Fibre Toxicol.*, 2015, **12**, 32.
78. P. M. Hinderliter, K. R. Minard, G. Orr, W. B. Chrisler, B. D. Thrall, J. G. Pounds and J. G. Teeguarden, ISDD: A computational model of particle sedimentation, diffusion and target cell dosimetry for in vitro toxicity studies, *Part. Fibre Toxicol.*, 2010, **7**.

79. J. T. Elliott, M. Rosslein, N. W. Song, B. Toman, A. Kinsner-Ovaskainen, R. Maniratanachote, M. L. Salit, E. J. Petersen, F. Sequeira, E. L. Romsos, S. J. Kim, J. Lee, N. R. von Moos, F. Rossi, C. Hirsch, H. F. Krug, W. Suchaoin and P. Wick, Toward Achieving Harmonization in a Nanocytotoxicity Assay Measurement Through an Interlaboratory Comparison Study, *Altex*, 2017, **34**, 201-218.
80. M. Rösslein, J. T. Elliott, M. Salit, E. J. Petersen, C. Hirsch, H. F. Krug and P. Wick, Use of Cause-and-Effect Analysis to Design a High-Quality Nanocytotoxicology Assay, *Chem. Res. Toxicol.*, 2015, **28**, 21-30.
81. G. J. Oostingh, E. Casals, P. Italiani, R. Colognato, R. Stritzinger, J. Ponti, T. Pfaller, Y. Kohl, D. Ooms, F. Favilli, H. Leppens, D. Lucchesi, F. Rossi, I. Nelissen, H. Thielecke, V. F. Puntes, A. Duschl and D. Boraschi, Problems and challenges in the development and validation of human cell-based assays to determine nanoparticle-induced immunomodulatory effects, *Part. Fibre Toxicol.*, 2011, **8**, 21.
82. Malvern Instruments Limited. Technical Note: Intensity - Volume - Number: Which size is correct, 2017.
83. F. Caputo, J. Clogston, L. Calzolai, M. Rösslein and A. Prina-Mello, Measuring particle size distribution of nanoparticle enabled medicinal products, the joint view of EUNCL and NCI-NCL. A step by step approach combining orthogonal measurements with increasing complexity, *J. Controlled Rel.*, 2019, **299**, 31-43.

Table 1 – Summary table listing techniques, mode of operation, laboratories that used this technique, etc.

Technique	Mode of operation	Number of laboratories that used this technique	Direct Measurement of NP Number Concentration	Does size measurement include coating?
SEM	Measures scattered electrons off of or through a sample	2	N	N
DLS	Measures Brownian motion of particles using a laser	1	N	Y
spICP-MS	Measures signal intensity for a given element for a single particle	2	Y	N
NTA	Uses light scattering and Brownian motion using a laser to measure the particle size distribution	1	Y	Y
DMA	Separates charged aerosilized particles according to their mobility in an electric field	1	N	Y

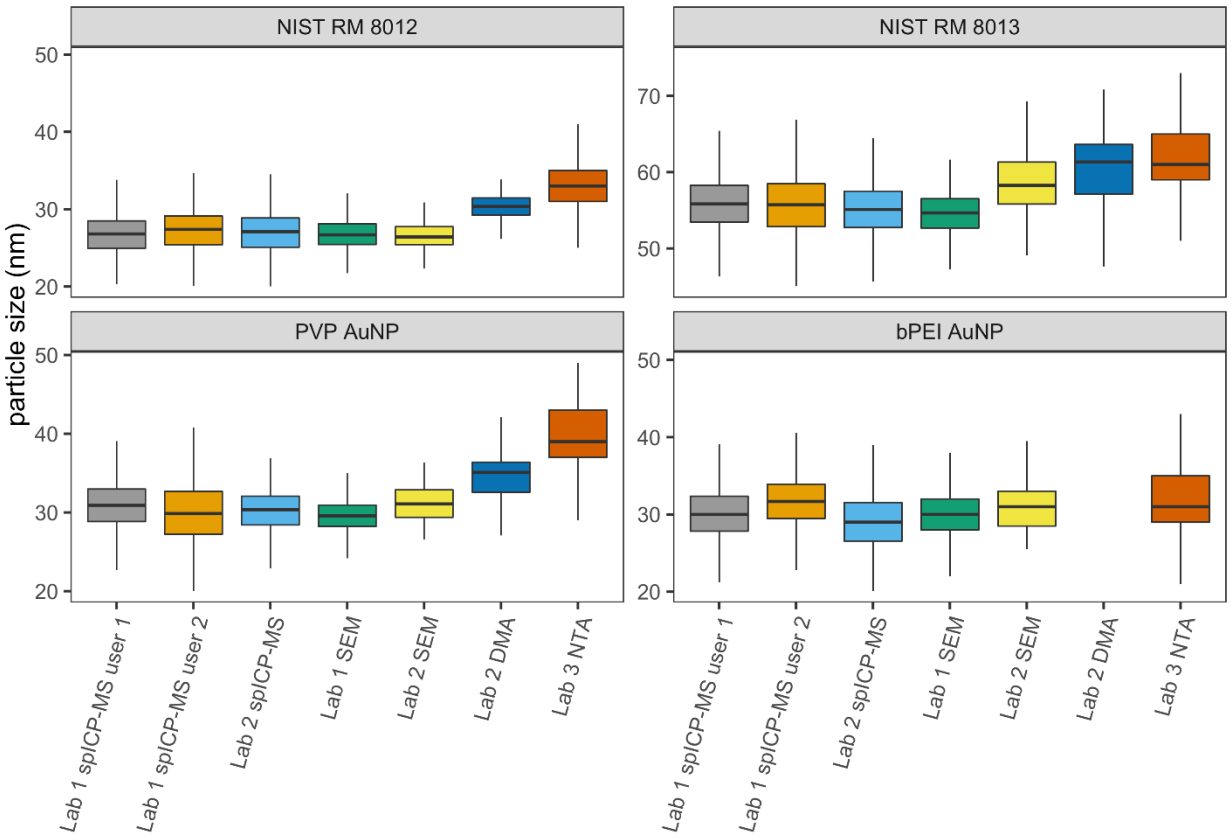


Figure 1 – Boxplots for the NIST RM 8012, NIST RM 8013, PVP AuNP, and bPEI AuNP (in plastic vials) samples. The thick horizontal line across each box marks the median of the corresponding particle size distribution, and the bottom and top of the box indicate the 25th and 75th percentiles, respectively. The bottom and top whiskers indicate the range for 10th and 90th percentiles, respectively. Values are provided for spICP-MS, DMA, NTA, and SEM. Given the broad size distribution of DLS relative to other techniques, boxplots including the DLS size distributions are provided in Figure S3. Results are not reported for the DMA analysis of the bPEI sample because of challenges with analyzing this sample.

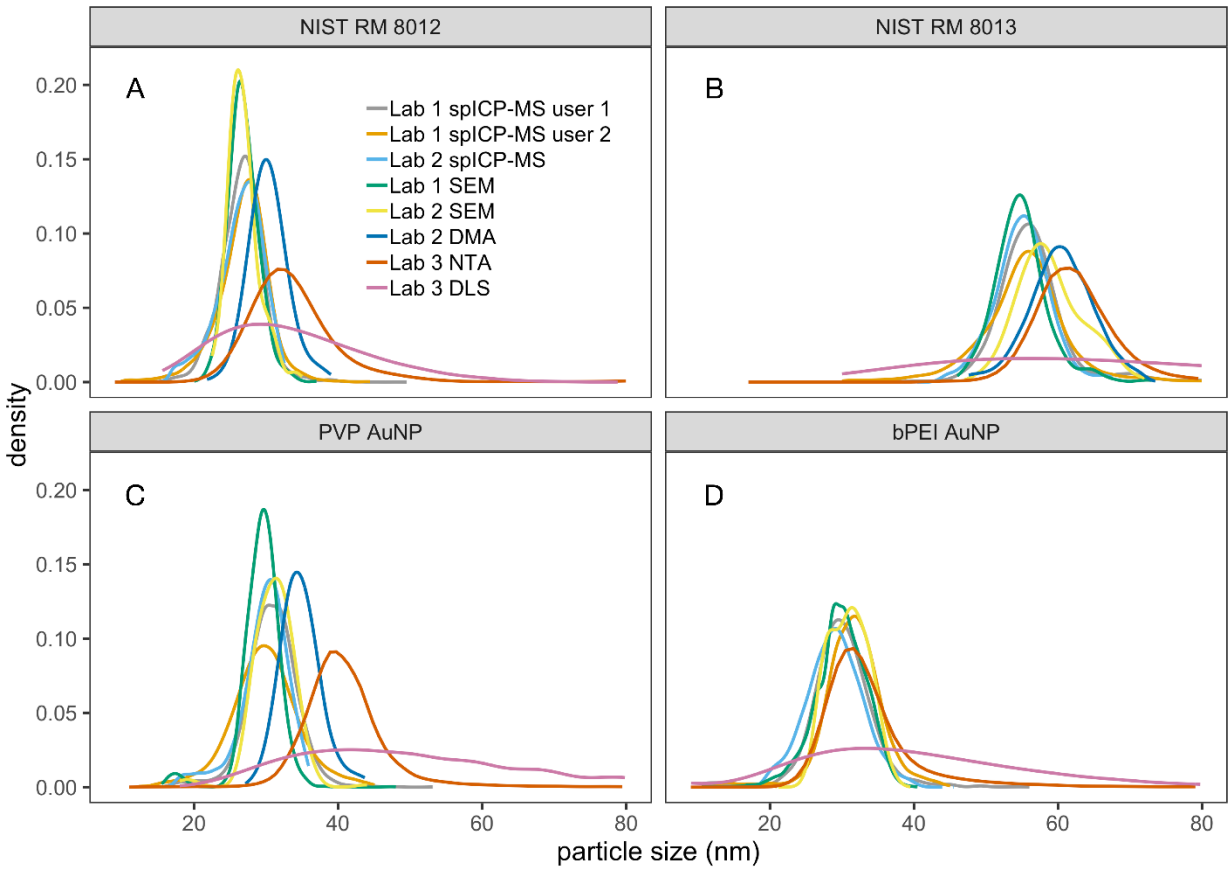


Figure 2 – Kernel density plots for NIST RM 8012 (A), NIST RM 8013 (B), PVP AuNP (C), and bPEI AuNP (in plastic vials) samples (D). Values are provided for spICP-MS, DMA, NTA, SEM, and DLS. Results are not reported for the DMA analysis of the bPEI sample because of challenges with analyzing this sample.

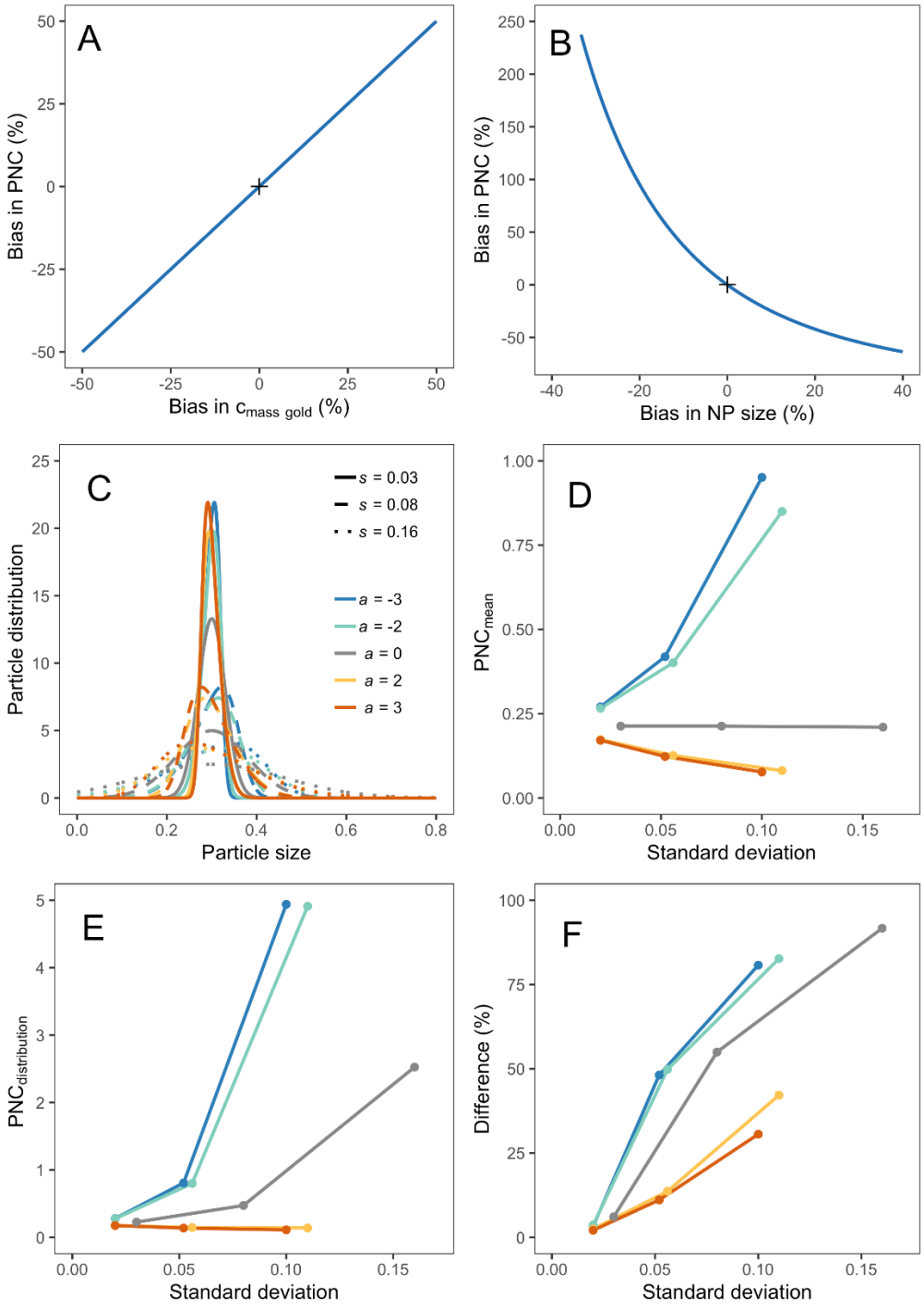


Figure 3 – Modeling for impact of size distribution changes on NP number concentration measurements. Plots show the bias in the derived PNC for a bias in the measured elemental concentration (A) or NP size (B). Fifteen different distributions were generated using a skew Normal distribution to model the impact of skew (*a*) and standard deviation of the distribution (*s*) (C). The impact of different amounts of skew or standard deviations of the distribution was evaluated for PNC_{mean} (D), $PNC_{distribution}$ (E), or the percentage different between $PNC_{distribution}$ and PNC_{mean} (calculated as 100 % * ($PNC_{distribution}$-PNC_{mean})/$PNC_{distribution}$) (F).

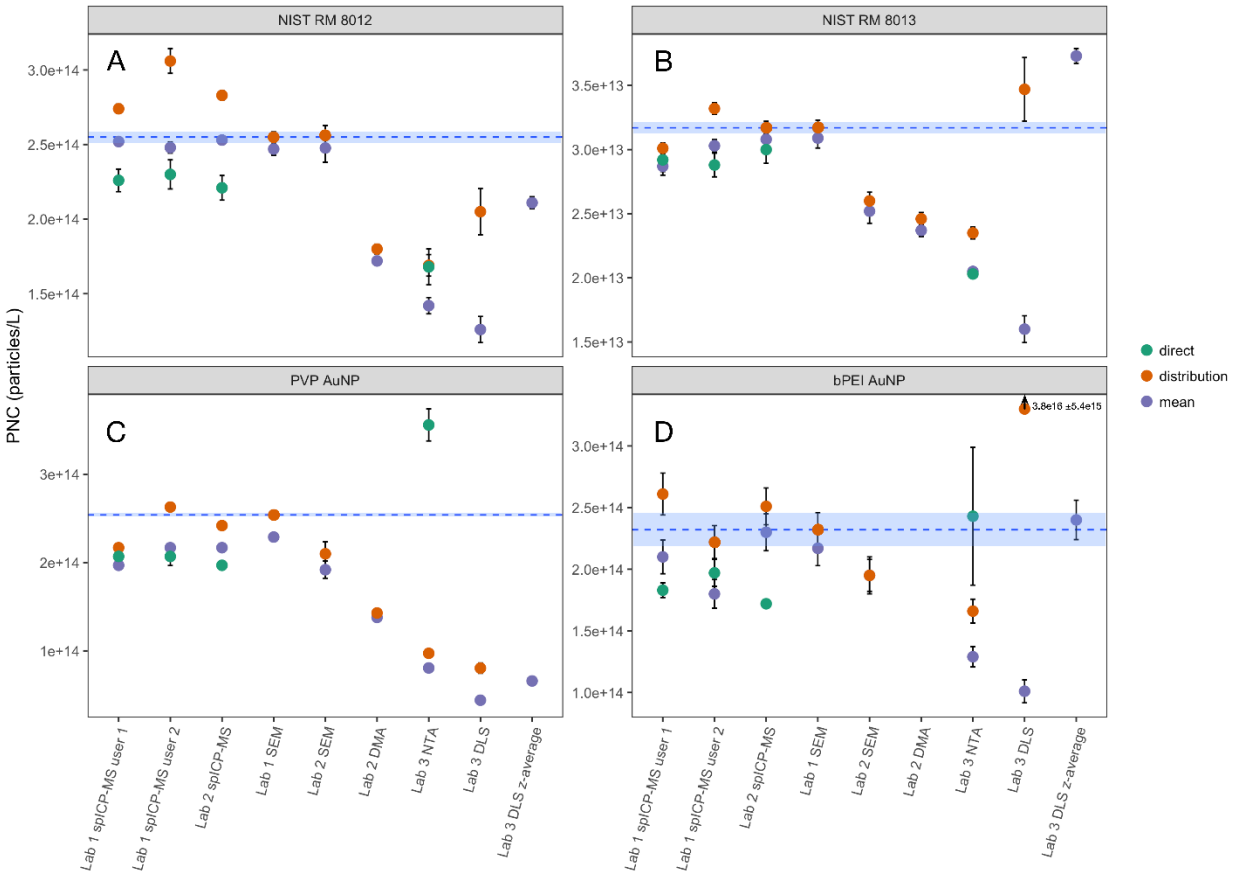


Figure 4 – Comparison among techniques, laboratories, and operators for the PNC measurements (PNC_{mean} (purple circles), $PNC_{distribution}$ (orange circles), or PNC_{direct} (green circles)) for (A) NIST RM 8012, (B) NIST RM 8013, (C) PVP AuNP, and (D) bPEI AuNP (in plastic vials) samples. Values are provided for spICP-MS, DMA, NTA, SEM, and DLS. Data points indicate the mean and the error bars are 95 % confidence intervals, and error bars that are not visible are smaller than data points. The horizontal dotted blue line and the blue shaded area correspond to the mean and 95 % confidence interval, respectively, of the $PNC_{distribution}$ results for SEM analyses from Laboratory 1.

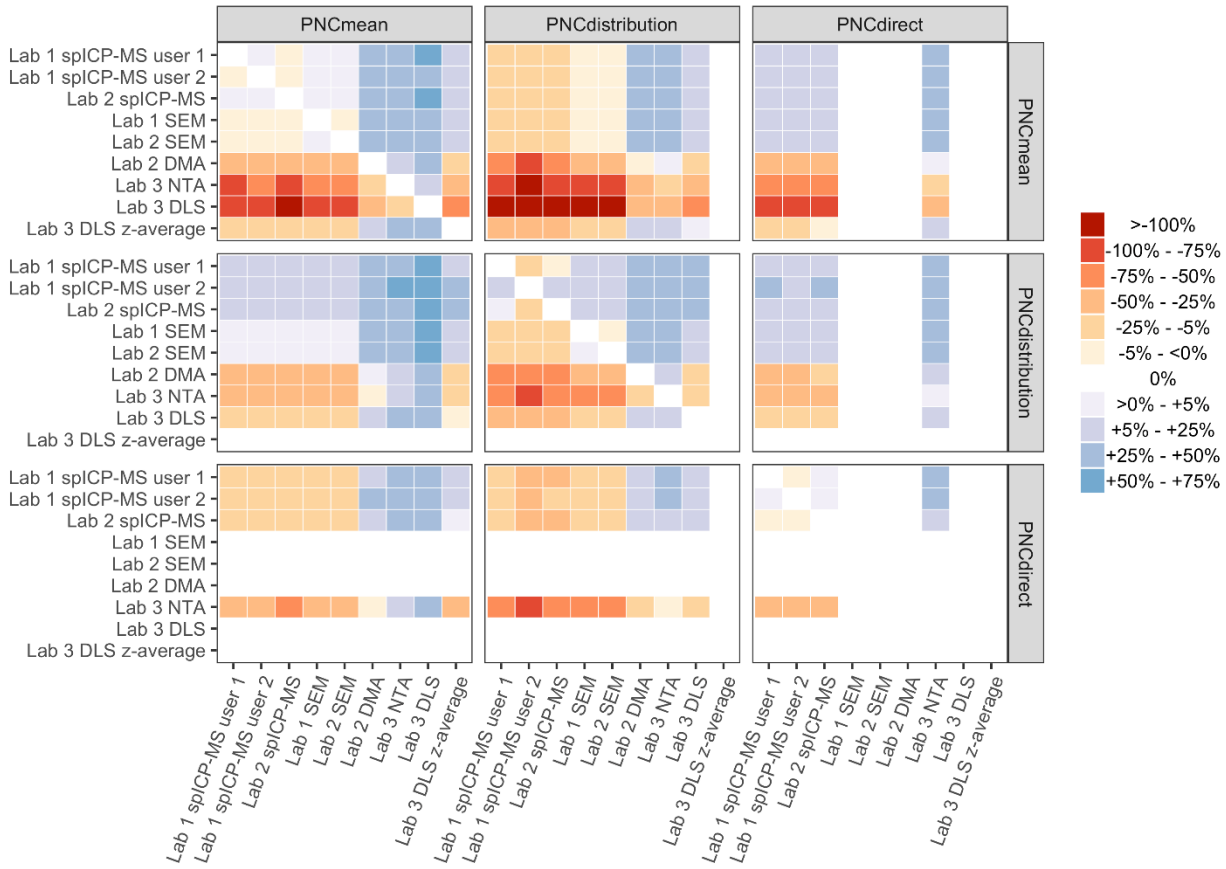
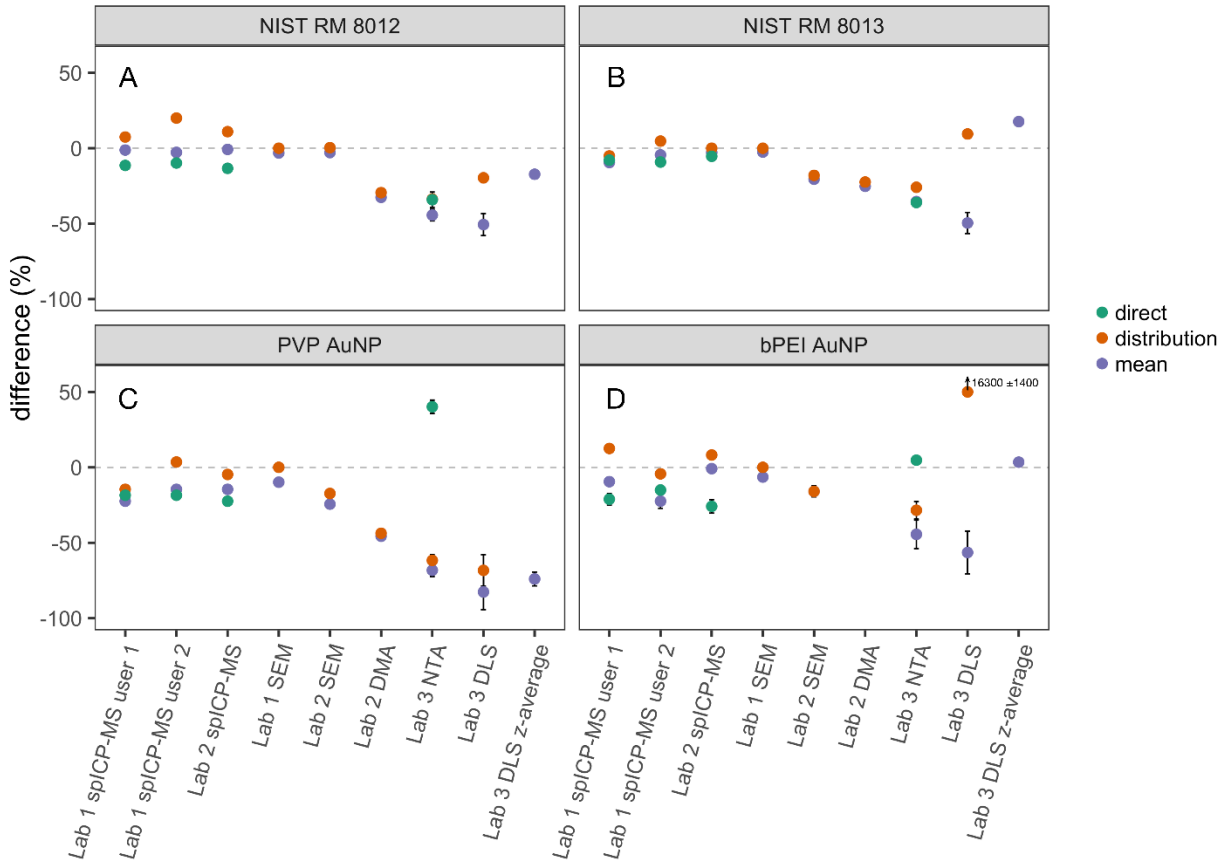


Figure 5 – Pairwise comparison among all techniques for the RM 8012 sample for the PNC_{mean}, $PNC_{distribution}$, and PNC_{direct} values. All values are percentages calculating using the formula 100 % * (PNC_y – PNC_x)/PNC_y where PNC_x is the PNC listed in the column and PNC_y is the PNC listed in the row. Colors indicate the percentage deviation between the techniques.



1073

1074 Figure 6 – Comparison of PNC_{direct}, PNC_{mean}, and $PNC_{distribution}$ to $PNC_{distribution}$ using SEM by
1075 Laboratory 1 for samples RM 8012 (A), RM 8013 (B), PVP AuNP (C), and bPEI (in plastic vials)
1076 AuNP (D). Data points indicate the mean and the error bars are the propagated errors for two times
1077 the relative uncertainty, and error bars that are not visible are smaller than data points. Data were
1078 calculated using the following formula:
1079 $Percentage=100\ \% * (PNC-PNC_{distribution,\ SEM})/PNC_{distribution,\ SEM}$.

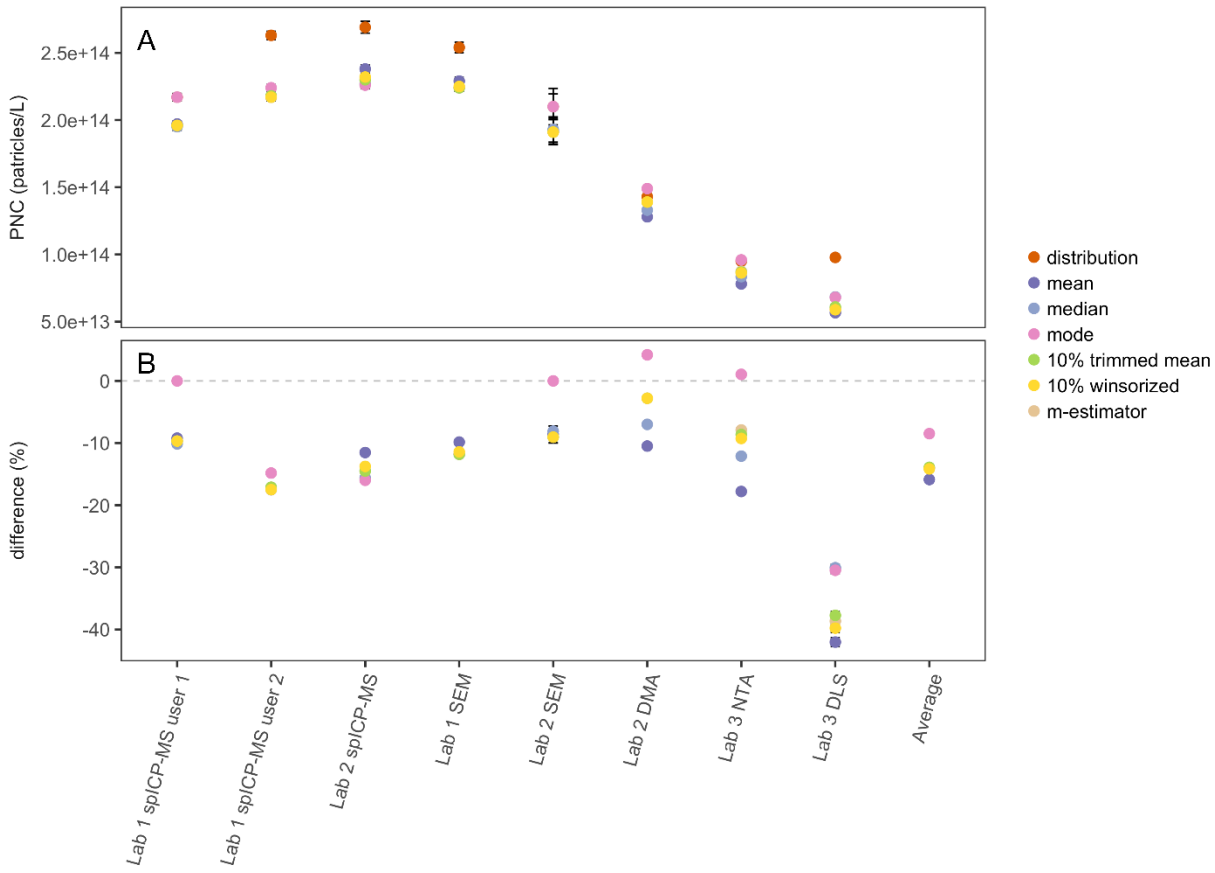


Figure 7 – Comparison of derived PNC values from a range of central tendency indicators and the particle size distribution for the PVP AuNPs. Data is shown for the derived PNC values (A) or by comparing the values for the different central tendency indicators against $PNC_{distribution}$ (B). Percentages were calculated using the formula 100 % * ($PNC_{distribution}$ – $PNC_{central\ tendency\ indicator}$)/$PNC_{distribution}$. Data indicate the mean value and error bars indicate 95 % confidence intervals, and error bars that are not visible are smaller than data points.

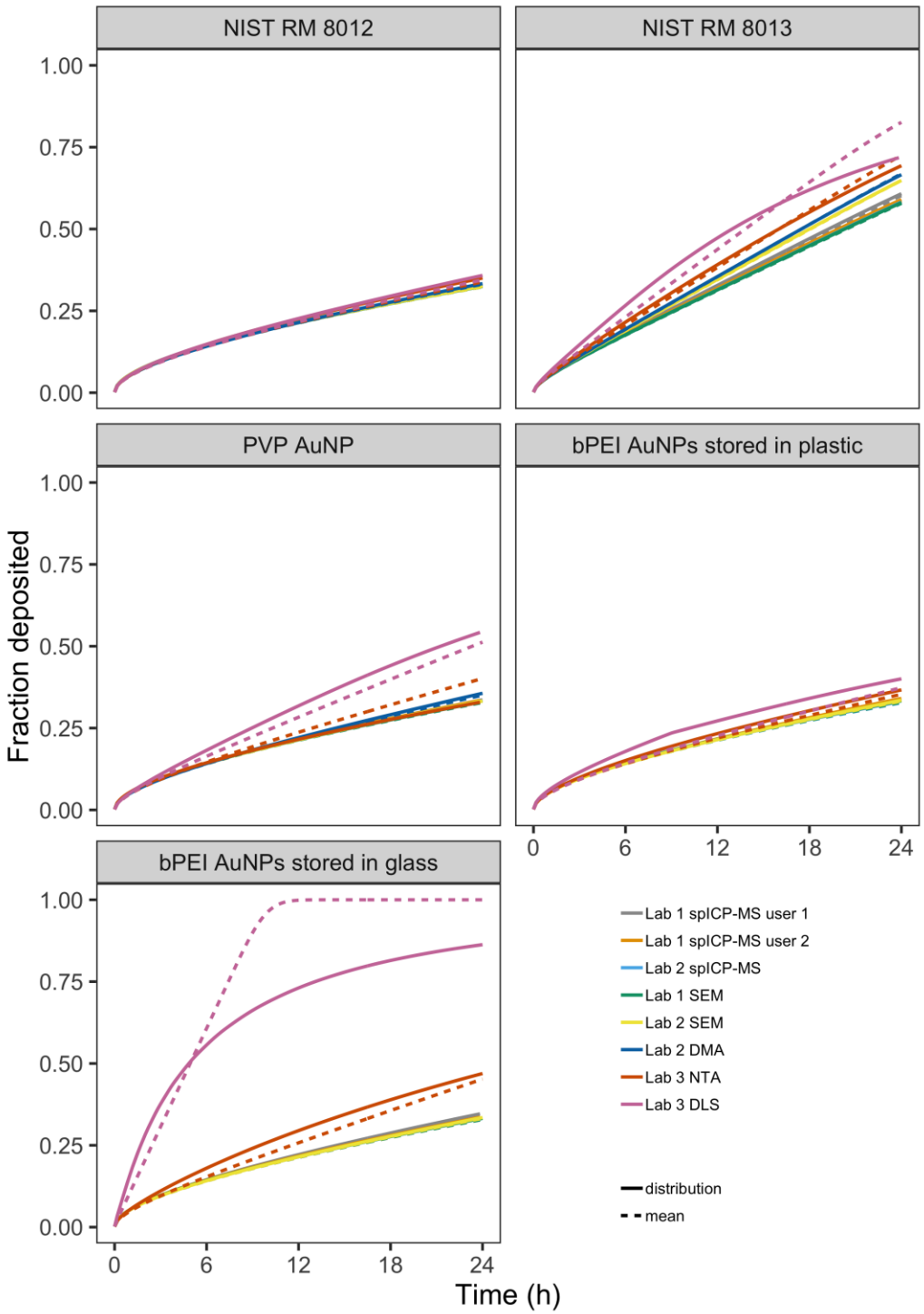


Figure 8 – Modeled *in vitro* concentrations of RM 8012, RM 8013, PVP AuNPs, and bPEI AuNPs stored in plastic or glass using the DG-ISDD model. This model does not provide an estimate of uncertainty for each data point and therefore uncertainty values are not included.