NanoImpact 21 (2021) 100277

Contents lists available at ScienceDirect

NanoImpact

journal homepage: www.elsevier.com/locate/nanoimpact

A meta-analysis framework to assess the role of units in describing nanoparticle toxicity

Robert M. Wheeler *, Steven K. Lower

The Ohio State University, Columbus, OH 43210, United States of America

ARTICLE INFO

Keywords:
Nanoparticle
Nanotoxicology
Meta-analysis

ABSTRACT

Despite ample research on nanoparticles, their environmental toxicity is still debatable. The lack of consensus is due in part to the challenge of comparing studies because of variability in parameters like test organism, test medium, and duration of experiment. However, the *unit* used to compare the toxicology of nanoparticles is one variable that experimentalists can control. Traditionally, mass per volume is the most common unit used to make comparisons, but there is growing evidence that alternative units such as surface area per volume or particles per volume may provide a better and more mechanistic measure of toxicity. Herein, we propose and test a meta-analytic framework to study the effect of units on nanotoxicology using data from the NanoE-Tox database, a freely available database containing 1518 toxicology values from 224 published articles of which 42 records met our basic inclusion criteria. These data were augmented with more recent data published over the past five years as archived by the Web of Science citation index. An additional 27 records from 1676 papers met the inclusion criteria and were also included in the analysis. The meta-analysis framework measures the degree of heterogeneity for each of three units (grams/L, particles/L, surface area/L) grouped by the type of test organism, particle chemistry, and manner in which a nanoparticle's size was measured (e.g., nominal particle size reported by the manufacturer vs. measurement of size for particles suspended in the liquid medium used in a subsequent toxicity experiment). The result of the meta-analysis reveals that surface area per volume reduces the heterogeneity in the Ag crustacean subgroup when nanoparticle size was measured in the test medium, and the ZnO crustacean subgroup when nanoparticle size was measured out the test medium and may therefore be a more appropriate estimate of the toxicity of soluble nanoparticles. No subgroups in our analysis showed a reduction in heterogeneity for particles per volume in either soluble or insoluble nanoparticles. The lack of conclusion on insoluble nanoparticles was *not* due to a limitation of our meta-analysis but rather highlights a critical deficiency in the primary literature. The majority of published studies fail to report common measures of error that are essential for further analysis (i.e. error of the measured nanoparticle size and/or interoperable error of the measured half-maximal concentration of the toxic endpoint). If future nanotoxicity studies report such error, as they should, then the framework of our meta-analysis could be used more broadly to provide a simple, statistically rigorous way to assess the role of units on the toxicity of nanoparticles.

1. Introduction

Nanomaterials are poised to be the defining technology of the next several decades with uses in batteries, drug delivery, photovoltaic cells, and wastewater treatment (Saunders and Turner 2008; Qu et al. 2013; Blanco et al. 2015; Caballero-Guzman and Nowack 2016; Chen et al. 2016). Despite potential applications, there is legitimate concern about the environmental and health implications of nanomaterials, particularly when these particles enter the environment after their intended use (Klaine et al. 2008; Elsaesser and Howard 2012; von der Kammer et al. 2012; Schrurs and Lison 2012; Bundschuh et al. 2016; Salieri et al. 2018). As evidence of this, 1676 articles on nanotoxicology have been published between January 2015 and April 2020 (Web of Science, key word search). Despite this research, there is little consensus as to the degree of harm that nanomaterials pose to humans or the environment (Alkilany et al. 2016; Steinhäuser and Sayre 2017; Arvidsson 2018; Bundschuh et al. 2018).

The lack of agreement is due in part to the fact that toxicity studies

Abbreviations: MPV, Mass Per Volume; PPV, Particles Per Volume; SAPV, Surface Area Per Volume; OTM, Out of Test Medium; ITM, In Test Medium.

* Corresponding author at: 275 Mendenhall Laboratory, 125 South Oval Mall, Columbus, OH 43210, United States of America.
E-mail addresses: wheeler.1010@osu.edu (R.M. Wheeler), lower.9@osu.edu (S.K. Lower).

https://doi.org/10.1016/j.impact.2020.100277
Received 1 September 2020; Received in revised form 13 November 2020; Accepted 16 November 2020
Available online 3 December 2020

use different test mediums, different test organisms, and various durations of exposure. Comparisons in nanotoxicology are more difficult because researchers also use different methods to characterize the nanoparticle itself (Reidy et al. 2013; Hua et al. 2016; Jemec et al. 2016; Keller et al. 2017). There is, however, one universal factor that can be controlled and could lead to greater consensus on nanotoxicity: the *unit* used to express the dose. Toxicity results are typically reported as mass per volume (MPV, g/L). This dose, while commonly used, may be biochemically inappropriate. MPV notation treats nanoparticles as a collection of atoms or molecules instead of a distinct chemical entity, and this may lead to erroneous comparisons as illustrated in Fig. 1.

Indeed, editors of toxicology journals as well as the International Organization for Economic Co-operation and Development (OECD 2017; Kraegeloh 2018) have recently called on researchers to determine the ideal way to dose nanoparticles. Studies on individual test organisms demonstrate the value of alternative dosing metrics like surface area for soluble metal particles and carbon nanoparticles or number of particles for insoluble metal particles (Kennedy et al. 2015; Mottier et al. 2016; Lagier et al. 2017). But, this work is largely limited to individual studies rather than systematic review, with the only other broader attempt for aquatic nanoparticles being the use of equi-response curves that show volume to be the best metric followed by surface area (Hua et al. 2016). Perhaps it is time to question the basic assumption that environmental toxicity should be measured as MPV instead of particles/L or surface area/L.

Herein, we sought to capture a wider survey of the literature to determine if the effect seen at the level of a single report is more universal. A meta-analysis was conducted to assess whether nanotoxicity measurements are affected by the choice of unit (mass, number of particles, or surface area per volume). We hypothesized that surface area per volume (SAPV) would be the ideal unit to describe the toxicity of soluble (e.g. Ag, ZnO, CuO) or otherwise surface active particles (e.g. TiO_2); whereas particles per volume (PPV) would be the ideal dose unit for insoluble particles (e.g., CeO_2). The optimal 'unit of choice' (e.g., MPV, SAPV, PPV) was identified based on a reduction in the dataset's heterogeneity as measured by the H and Q statistics, which are standard, unitless heterogeneity metrics in meta-analysis.

Our meta-analysis utilized a large number of studies captured by the NanoE-Tox database (Juganson et al. 2015) plus papers from an additional search of Web of Science articles published in the past five years.

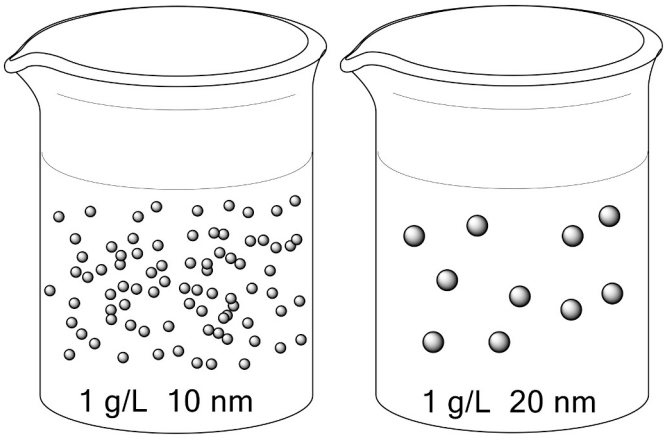

Fig. 1. Schematic drawing demonstrating the relationship between concentration (1 g/L) and size (10 vs. 20 nm particles). While both beakers have the same mass of Ag per volume, the 10 nm beaker has eight times as many particles as the 20 nm beaker and two times the total surface area. The effect seen here is somewhat extreme due to the sizes chosen, but a systematic error is introduced no matter the size because radii and mass concentration determine the derived units' concentrations. As a more general rule when comparing two particles of the same chemistry, the relative particle number scales as a function of r_2^3/r_1^3 and total surface area scales as a function of r_2/r_1.

A clear signal was observed for Ag and ZnO nanoparticles tested on crustaceans. For these soluble nanoparticles, our analysis revealed that SAPV was the best predictor of nanotoxicity. No subgroups in our analysis showed a preference for PPV. This was due in part to the stringent inclusion criteria used in our meta-analysis. Non-soluble particle chemistries expected to show a preference for PPV were "filtered out" because the source publications failed to report error in size and concentration. Nonetheless, our meta-analysis provides a statistically rigorous description and simple framework for assessing the role of units in nanotoxicity. As more studies are published with acceptable reporting of error, this approach will allow robust consideration of whether non-mass per volume units should be considered, which is a valuable asset for the design of studies in environmental nanotoxicity. Furthermore, this meta-analysis approach may be transferrable to other environmental health fields to evaluate, for example, whether surface area is the most relevant dose metric for acute (or chronic) nanoparticle lung toxicity, which has already been suggested in a retrospective analysis of animal studies (Schmid and Stoeger 2016).

2. Methods

Data for the meta-analysis was initially drawn from the NanoE-Tox database, which contains 1518 individual records where a record is defined as a row in the database (Juganson et al. 2015). Each record therefore represents one reported measurement of nanoparticle toxicity from one of the 224 different studies in the database. Since the NanoE-Tox database was last updated in 2015, the same NanoE-Tox search criteria was repeated using Web of Science on May 1, 2020 (minus 'carbon nanoparticle' as it was not of interest in the current meta-analysis). This yielded 1676 papers which were manually filtered, along with the NanoE-Tox studies, according to the inclusion criteria outlined below.

The choice to use combined datasets is not without its drawbacks as choices are made during data entry that may be different and/or incompatible with the choices made by Juganson et al. (2015). For example, in the 1676 reviewed papers published since 2015, several authors provided multiple characterizations of the nanoparticle in the test medium at several time points during the experiment. When faced with the choice of which characterization to use, we chose size measured at the earliest time point and favored size determined with transmission electron microscopy (TEM) over the dynamic light scattering when both were reported.

Data were processed using two different sets of inclusion criteria. The first, called the ***summary effect case,*** used strict inclusion criteria to calculate a more rigorous mean value of the dataset through error propagation. Four requirements had to be met: (i) metallic nanoparticle toxicity measurement in aquatic media; (ii) report the lethal concentration (LC_{50}) or effective concentration (EC_{50}) with mortality descriptor, namely, mortality, immobility, or luminesce and provide standard deviation of the toxic concentration; (iii) report the size and standard deviation of the nanoparticle; and (iv) particle shape reported as spherical or approximately spherical. Data that met these specifications were converted to the derived units, PPV expressed in particles/L and SAPV expressed in m^2/L, using **Eqs. (1)** and **(3)** respectively, shown in Table 1. The error in concentration and size was propagated according to in **Eqs. (2)** and **(4)** of Table 1.

The second set of criteria, called the ***dispersion effect case*** analysis, was created with three purposes: (i) to assess the effect of the four inclusion criteria on the summary effect case means, (ii) to increase data for the analysis, and (iii) to assess the effect of the units based on a change in heterogeneity, which is not possible due to error propagation in the summary effect case. To create the dispersion effect case, we relaxed the last two criteria (i.e., reported error in size and spherical shape) of the summary effect case analysis. The dispersion effect case still requires that nanoparticle size be provided but because error is no longer propagated, the reported size error is not required. Therefore, in

Table 1
Equations used to calculate the derived unit doses of particles per liter and surface area per liter for the summary effects analysis and dispersion effects analysis as well as the formulas used to propagate error in the dispersion effects analysis. ρ_{metal} is the density of the particle, $r_{particle}$ is the particle's radius, C and σ are the concentration and standard deviation of the concentration in g/L, particle/L or SA/L.

Equation #	Equation
1	$C_{particle}\left(\frac{particles}{L_{H_2O}}\right) = \frac{mass_{metal}}{L_{H_2O}} * \frac{1}{\rho_{metal}} * \frac{particle}{\left(\frac{4}{3}\pi (r_{particle})^3\right)}$
2	$\sigma_{particle} = C_{particle} * \sqrt{\left(\frac{\sigma_{mass}}{C_{mass}}\right)^2 + \left(3\frac{\sigma_{radius}}{r_{part}}\right)^2}$
3	$C_{Surface\ Area}\left(\frac{m^2}{L_{H_2O}}\right) = \frac{mass_{metal}}{L_{H_2O}} * \frac{1}{\rho_{metal}} * \frac{3}{r_{particle}}$
4	$\sigma_{Surface\ Area} = C_{Surface\ Area} * \sqrt{\left(\frac{\sigma_{mass}}{C_{mass}}\right)^2 + \left(\frac{\sigma_{radius}}{r_{part}}\right)^2}$

the case of the dispersion effect case, data were weighted based solely on concentration error converted to the derived units. Additionally, shape descriptors such as 'not applicable' (N/A in the NanoE-Tox database) and irregular were included in the dispersion effects analysis. MPV was converted to SAPV and PPV using **Eqs. (1)** and **(3)** in Table 1 as above with the assumption that all particles were spherical, and the reported size corresponded to diameter. For SAPV and PPV, the relative standard deviation of MPV was used to calculate the derived units' standard deviation.

Data from the NanoE-Tox database was filtered according to the inclusion criteria using MATLAB2017a while the literature search was filtered as it was manually entered. Both datasets had their units converted in MATLAB2017a and were subsequently exported to R (version 3.4.3) for further analysis. The meta-analysis utilized the 'metafor' (version 2.0–0) package to calculate the mean effects and the heterogeneity statistics using the restricted maximum likelihood method under the random effects assumptions which presumes each study is drawn from a distribution of effects that share the same mean, resulting in a basic accounting for the differences between studies. The 'metafor' package was also used to run the meta-regression calculations using the moderators, organic coating, and time of exposure (Viechtbauer 2010). The regression models were assessed based on whether Akaike Information Criteria corrected for small sample size (AICc) and Bayesian Information Criteria (BIC), relative measures of the quality of the statistical model, decreased compared to the simple mean values and whether a high R^2 value indicated a suitable fit.

The data were broken up into subgroups based on three qualifiers: the particle chemistry used in the experiment (e.g., Ag, CuO, etc.), the organism type (e.g., crustaceans, bacteria, etc.) used in the test, and whether the size of the nanoparticle was measured in the test medium (ITM) or out of the test medium (OTM). The definitions of ITM vs. OTM are based on a distinction made in the NanoE-Tox database. ITM refers to the size of the particle as measured in the test medium (e.g., liquid buffer) at the beginning of an experiment before addition of the test organism. This size was generally measured by dynamic light scattering or nanoparticle tracking analysis. OTM size is a bit more nebulous because it refers to the size of the material before it is placed in the test medium. This may be the nominal size listed by the manufacturer, size determined by TEM, or in a few instances the size of a particle in ultrapure water as determined by light scattering. The analysis was run for all subgroups that contained more than four records in the subgroup.

3. Results

3.1. Summary effect case

3.1.1. Description of data included in analysis

The summary effect case study was conducted by applying the following inclusion criteria to peer-reviewed publications: metallic nanoparticle, LC_{50} or EC_{50} values reported with standard deviation, nanoparticle size reported with standard deviation, and spherical particle shape (see **Methods**). After applying this relatively stringent criteria, only 42 records from four separate publications remained from the NanoE-Tox dataset. An additional 27 records from eight separate publications were added from a manually curated data set generated by using Web of Science to search for terms from the original NanoE-Tox database.

These records were then divided into subgroups based on the type of organism tested, the chemistry of the particle tested, and whether the size of the particle was measured in the test medium (ITM) or out of the medium (OTM). ITM refers to measurement of the size of particles that were suspended in the same liquid media that was used in the toxicity experiment; whereas OTM generally refers to the nominal size provided by the manufacturer of the particle. A total of 14 subgroups were generated with these summary effect inclusion criteria, namely: Ag/crustacean/OTM, Ag/zebrafish/OTM, CuO/crustacean/OTM, CuO/rotifer/OTM, TiO_2/algae/OTM, TiO_2/crustacean/OTM, ZnO/crustacean/OTM, Ag/bacteria/ITM, Ag/crustacean/ITM, CuO/crustacean/ITM, CuO/rotifer/ITM, TiO_2/crustacean/ITM, ZnO/bacteria/ITM, and ZnO/crustacean /ITM.

Of the 14 subgroups, the following seven contained greater than four records, a requirement for inclusion in the analysis: Ag/crustacean/OTM, Ag/zebrafish/OTM, CuO/crustacean/OTM, TiO_2/algae/OTM, Ag/bacteria/ITM, Ag/crustacean/ITM, and ZnO/bacteria/ITM (see Table 2). Of the seven excluded groups, five had only one record in the dataset, while two subgroups had two (CuO/crustacean/ITM) or three (ZnO/Crustacean/OTM) records.

3.1.2. Calculated means

Numerical results of the summary effects analysis are shown in Table 2**.** These results are also visually summarized as forest plots in Fig. 2 and **Fig. S1**. Forest plots show the mean and calculated 95% confidence interval for each record as well as the meta-analytic mean and its 95% confidence interval. Overall, Fig. 2 demonstrates that the records across all the units (MPV, PPV, SAPV) are relatively well clustered on a logarithmic scale with the visual clustering increasing in the derived units. **Fig. S1** shows a similar trend though to a lesser degree due to a smaller number of records in the subgroup and many of the subgroups are dominated by a single study.

Despite the clustering, when looking down a column in Fig. 2 (or **Fig. S1**), several subgroups contain at least one consistent outlier. We initially thought that the meta-regression could account for these outliers by using moderators of time or organic coating. However, this analysis was not particularly informative for summary effects because there was insufficient variability in most subgroups with respect to time or coating. For example, **Table S1** shows TiO_2/algae/OTM is dependent upon exposure time, as are ZnO/bacteria/ITM and Ag/bacteria/ITM. But this yields no new insight as all records in each subgroup were drawn from a single study that moderated time but not the coating (Mallevre et al. 2014; Ozkaleli and Erdem 2018). Similarly, a strong relationship with coating was found for the Ag/zebrafish/OTM subgroup. But closer examination of raw data revealed only two coatings, one from each paper, in the dataset indicating the relationship could just as easily be explained by other study-level differences. Despite the strong R^2 for the time moderator of the Ag/zebrafish/OTM subgroup, it is likely spurious as is the coating moderator for the Ag/crustacean/ITM subgroup, and both time and coating for the CuO/crustacean/OTM subgroup. This inference is supported by the increasing AICc value since it is designed to penalize statistical model fit to small datasets, which is likely occurring in each case.

The only subgroup with sufficiently diverse data to exhibit a relationship with either variable (time or coating) is Ag/crustacean/OTM. In this case, the low R^2 value indicates that time is a poor moderating factor for this subgroup, and coating (despite the high R^2 value) has an

Table 2
Mean LC_{50} or EC_{50} toxic effects in log units as calculated by the meta-analysis on the summary effects dataset broken out by (i) whether the size of the nanoparticle was measured *in* or *out* of the liquid medium used in the toxicology test, (ii) the chemistry of the particle, and (iii) the organism type included in the study. The columns report the meta-analytic mean and standard error, the prediction interval is the expected range where 95% of future results will fall, T^2 is the between studies variance, H and Q are relative measures of heterogeneity. Also cited are the studies included in the analysis and the number of records in the category (*n*). g/L = grams per liter; part/L = particles per liter; SA/L = surface area per liter in units of m^2/L.

			Mean	Standard Error	Prediction Interval	T^2 Estimate	H	Q	Studies Included in Analysis
Size Measured Outside of the Test Medium	Ag Zebrafish (*n* = 7)	g/L	−4.36	0.08	0.59	0.21	81.54	474	Ribeiro et al. (2014); Boehme et al. (2015)
		part/L	12.19	0.32	1.99	0.70	3.40	28	
		SA/L	−2.85	0.15	1.02	0.37	5.96	46	
	Ag Crustacean (*n* = 31)	g/L	−4.45	0.10	1.19	0.57	727.04	18,611	Blinova et al. (2013); Ribeiro et al. (2014); Silva et al. (2014); Ulm et al. (2015); Borase et al. (2019)
		part/L	12.56	0.19	2.03	0.97	8.22	248	
		SA/L	−2.77	0.12	1.31	0.63	14.58	353	
	TiO_2 Algae (*n* = 8)	g/L	−1.81	0.18	1.32	0.51	5843.42	9574	Ozkaleli and Erdem (2018)
		part/L	14.02	0.18	1.22	0.46	5.57	39	
		SA/L	−0.53	0.18	1.31	0.50	41.42	265	
	CuO Crustacean (*n* = 4)	g/L	−2.35	0.49	4.75	0.99	854.35	1071	Kim et al. (2017); Rotini et al. (2018)
		part/L	12.52	0.24	1.47	0.25	1.36	3	
		SA/L	−1.36	0.35	3.23	0.67	22.50	60	
Size Measured Inside of the Test Medium	ZnO bacteria (*n* = 5)	g/L	−1.07	0.04	0.29	0.08	183.64	281	Mallevre et al. (2014)
		part/L	12.72	0.04	0.28	0.08	6.68	20	
		SA/L	−0.29	0.04	0.29	0.08	38.27	66	
	Ag bacteria (*n* = 12)	g/L	−2.12	0.11	0.90	0.39	222.22	501	Mallevre et al. (2014)
		part/L	12.52	0.28	2.21	0.95	187.12	2259	
		SA/L	−1.23	0.14	1.15	0.50	145.02	1324	
	Ag Crustacean (*n* = 4)	g/L	−5.53	0.31	2.96	0.61	775.53	782	Silva et al. (2014); Ulm et al. (2015)
		part/L	11.27	0.31	2.74	0.56	18.71	39	
		SA/L	−3.98	0.24	2.25	0.46	59.07	80	

inconsistent signal for AICc and BIC in the derived units, though it does have a strong signal in the MPV case (see **Table S1**). Despite statistics showing some relationship with the proposed moderating variables, the outliers could also be explained as small datasets from different studies (e.g., see Fig. 2 Ag/crustacean/OTM**)**.

The means of the summary effect analysis represent a statistically rigorous estimate of the toxicity of nanoparticles because of the unbiased inclusion criteria and strict accounting of error in the measured nanoparticle concentration. Despite this rigor, the summary effects analysis cannot be used to assess the hypothesis that the derived units of PPV or SAPV are better than MPV because the size error is propagated. The two chosen statistics, H and Q (described in detail in the **Supplemental Information**) used to assess heterogeneity in the analysis, both include a term normalized by the study level variation. This normalization allows for comparison across units, but due to the propagation of the size error, the study level error increases greatly for derived units, which in turn decreases heterogeneity. Therefore, any observed decrease in H or Q cannot be attributed solely to changing units. Instead, it is attributed to the increase in the study level variation.

The TiO_2/algae/OTM column in Fig. 2 highlights the effect of error propagation nicely. In the MPV panel there are eight datapoints with little to no overlap in the 95% confidence intervals. In the SAPV case, where, as Table 1 shows, error is propagated by adding the size error to the concentration error, each record's 95% confidence interval begins to overlap. For PPV, where error in size is multiplied by nine before being added to the concentration error, the 95% confidence intervals almost all overlap. This visual change is borne out by the heterogeneity statistics where H drops from 5843.42 in the MPV case to 41.42 in the SAPV case to 5.57 in the PPV case. Q shows a similar drop from 9574 in the MPV case to 265 and 39 for SAPV and PPV, respectively. Therefore, to assess the effect of units, the dispersion effect analysis must be used. The dispersion effect does not propagate size error thereby producing a less accurate mean but clearer picture of the role of units in the data's heterogeneity.

3.2. Dispersion effect case

Loosening the restrictions of the summary effect analysis successfully increased the number of records to 114 unique measurements from 16 different studies sorted into 18 different subgroups. Of the 18 subgroups, 10 had more than four records required for calculation of the mean effect. These results are reported in Table 3 and visually summarized in Fig. 3. Forest plots for all subgroups and units can be found in Fig. 4 and **Figs. S2** and **S3**. Surprisingly, the relaxation of the restrictions did not increase the number of records for four of the seven subgroups from the summary effect analysis. Only CuO/crustacean/OTM, Ag/crustacean/ITM and Ag/bacteria/ITM exhibited an increase in records by 16, 20 and 6 records respectively. This illustrates the effect of increasing data when error is not propagated on the final result as well as the effect of the lack of proper error propagation.

3.2.1. Comparing dispersion and summary effect analysis

Fig. 5 shows the mean values and prediction interval for each subgroup (with more than 4 records) for the summary versus the dispersion subgroups. For the four subgroups that exhibited no increase in records (Ag/zebrafish/OTM, Ag/crustacean/OTM, TiO_2/algae/OTM, and ZnO/bacteria/ITM), the means for the MPV units were the same, as expected because there was no change in weighting. The mean values for the PPV and SAPV also remained constant across the two case analyses for the subgroups with no record changes.

This goes against expectations from a theoretical perspective, as one would expect each record's weight to change when error was not propagated in the dispersion effects and therefore the mean to change in response. For the TiO_2/algae/OTM and ZnO/bacteria/ITM subgroups the lack of change is attributed to all of the records having the same size assigned to them and therefore the relative weighting remains the same.

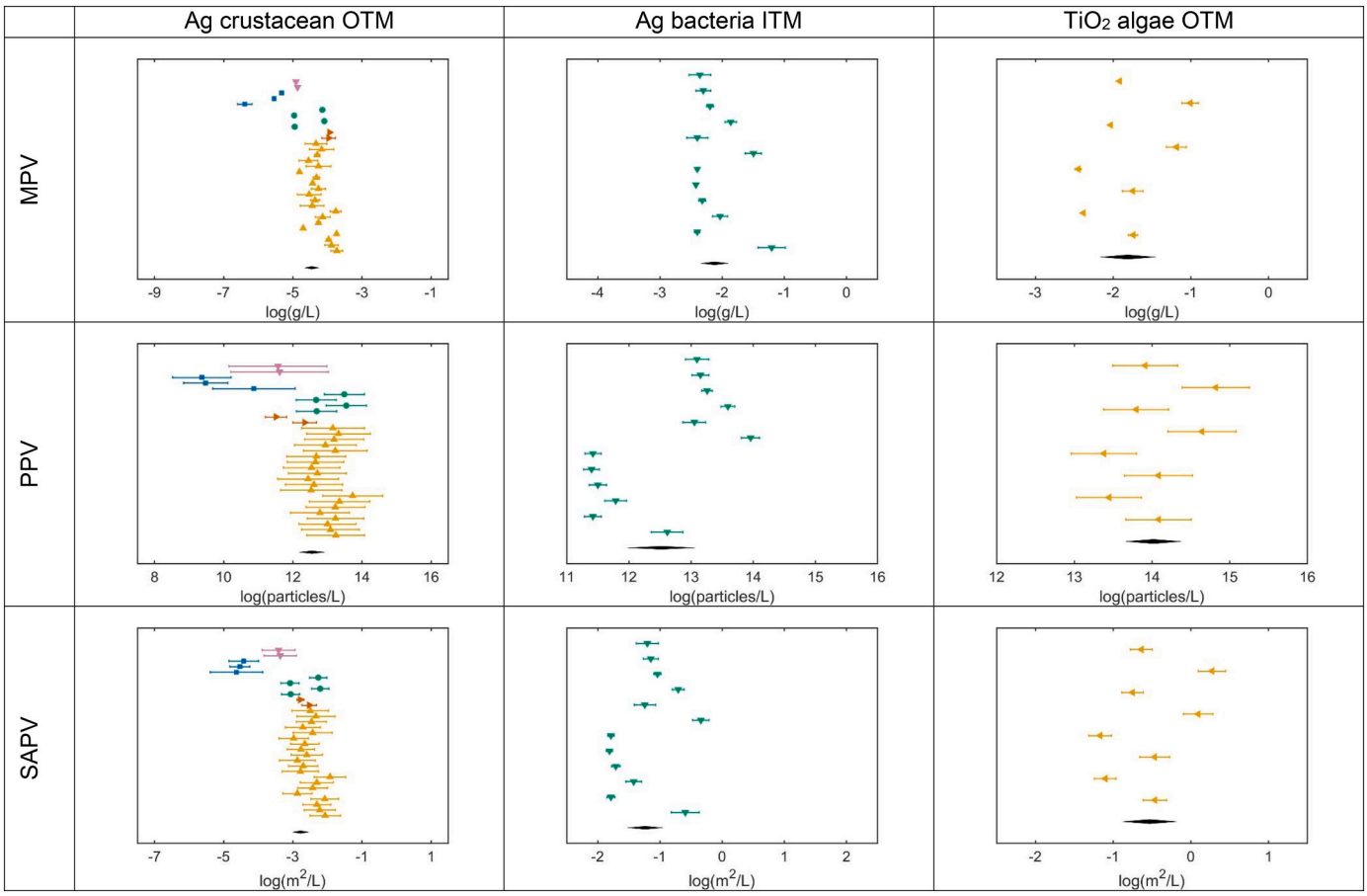

Fig. 2. Forest plots of the summary effects analysis for each combination of particle, organism, and chemistry as defined by the three units: grams per volume (MPV), particles per volume (PPV), and surface area per volume (SAPV). OTM = the size of the nanoparticle was measured out of the test medium. ITM = nanoparticle size determined in the test medium. Each point represents the mean and the calculated 95% CI for each record. The black diamond (located at the bottom of each panel) is centered on the mean value from the meta-analysis where the width is the 95% CI values for the mean. The different colors in the silver crustacean column represent the data source. The Ag Zebrafish OTM column contains records from Blinova et al. (2013) represented by yellow upward-pointing triangles, Ribeiro et al. (2014) represnted by green circles, Silva et al. (2014) represented by blue squares, Ulm et al. (2015) represented by pink downward-pointing triangles, Borase et al. (2019) represented by orange right-pointing triangles. All data in the Ag-bacteria column come from Mallevre et al. (2014) represented by green downward-pointing triangles. All data in the TiO_2-algae-OTM column comes from Ozkaleli and Erdem (2018) represented by yellow left-pointing triangles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A similar explanation is attributed to the Ag/zebrafish/OTM subgroup, where only two studies were represented, but six of the seven records came from the same study causing a small change in the mean and prediction interval. For the Ag/crustacean/OTM data, where five studies are represented, the lack of change is odd but there is some slight variation indicating that weighting of each record changed but not enough to greatly alter the mean value (see Table 3).

For the three subgroups that exhibited a change in the number of records (CuO/crustacean/OTM, Ag/crustacean/ITM and Ag/bacteria/ITM), the means of the dispersion effect were within the prediction interval of the summary effects. Furthermore, the additional data served to reduce the prediction intervals of both the MPV mean and the derived units. The one exception to this trend is the PPV case for CuO/crustacean/OTM, which had an increase in prediction interval, and the dispersion effects mean fell outside of the boundaries of the prediction interval. While the derived units in the dispersion effects case appears to have greater precision than the summary effects case, the dispersion effects case is less accurate. This is confirmed by an observed increase in heterogeneity indicated by the H statistic.

Conversely, the MPV means are more accurate than their summary effects counterparts because both H and prediction interval decreased with the added data. We do not consider Q for the data with increased number of records because Q is highly dependent on the number of studies, a problem the H statistic was developed to solve (Higgins and Thompson 2002). Another indication that the derived unit means in the dispersion case are less precise comes from the dispersion effects case where no data was added and both measures of heterogeneity, H and Q show at least an order of magnitude increase compared to the summary effects case analysis for the derived units (Table 3). Of course, this increase was expected because error propagation artificially deflated the heterogeneity.

3.2.2. *Assessing heterogeneity differences between units*

Our hypothesis, mentioned in the Introduction, predicts that there will be a decrease in heterogeneity for SAPV with respect to soluble and surface-active particles; whereas PPV will show a decrease in heterogeneity for non-soluble particles. This analysis was limited to subgroups with more than eight records, the minimum number of records recommended for a reliable computation of the H statistic (Higgins and Thompson 2002). Seven subgroups met these criteria: ZnO/crustacean/OTM, Ag/crustacean/OTM, Ag/bacteria/OTM, TiO_2/algae/OTM, CuO/crustacean/OTM, Ag/crustacean/ITM and Ag/bacteria/ITM.

As seen in Table 3, Ag/crustacean/ITM and ZnO/crustacean/OTM showed a reduction in heterogeneity for the SAPV by 39.2 and 6.1% in H and 24.5% and 2.2% in Q respectively. For the other subgroups, two failed to show any signal (Ag/bacteria/OTM, TiO_2/algae/OTM). The

Table 3
Mean LC_{50} or EC_{50} toxic effects in log units as calculated by the meta-analysis on the dispersion effects dataset broken out by (i) whether the size of the nanoparticle was measured *in* or *out* of the liquid medium used in the toxicology test, (ii) the chemistry of the particle, and (iii) the organism type included in the study. The columns report the meta-analytic mean and standard error, the prediction interval is the expected range where 95% of future results will fall, T^2 is the between studies variance. Also shown are the relative measures of heterogeneity (H and Q) as well as their percent change for the mass per volume (g/L) case. The final column cites the studies included in the analysis, where *n*, shown in the first column, is the number of records in the category. g/L = grams per liter; part/L = particles per liter; SA/L = surface area per liter in units of m^2/L; N/A = not applicable (undefined).

			Mean	Standard Error	Prediction Interval	T^2 estimate	H	Q	H Percent Change	Q Percent Change	Studies Included in Analysis
Size Measured Outside of the Test Medium	ZnO bacteria ($n = 5$)	g/L	−1.07	0.04	0.29	0.08	13.55	281	N/A	N/A	Mallevre et al. (2014)
		part/L	14.59	0.04	0.29	0.08	13.55	281	0	0	
		SA/L	0.34	0.04	0.29	0.08	13.55	281	0	0	
	ZnO crustacean ($n = 17$)	g/L	−2.67	0.10	0.89	0.41	7.47	1291	N/A	N/A	Blinova et al. (2010); Kim et al. (2017)
		part/L	12.54	0.13	1.20	0.54	9.99	1282	−0.7	33.7	
		SA/L	−1.44	0.10	0.87	0.40	7.31	1212	−6.1	−2.2	
	Ag bacteria ($n = 18$)	g/L	−2.08	0.08	0.74	0.34	10.83	554	N/A	N/A	Mallevre et al. (2014); Mallevre et al. (2016)
		part/L	14.65	0.08	0.74	0.34	10.83	554	0	0	
		SA/L	−0.50	0.08	0.74	0.34	10.83	554	0	0	
	Ag zebrafish ($n = 7$)	g/L	−4.36	0.08	0.59	0.21	9.03	474	N/A	N/A	Ribeiro et al. (2014); Boehme et al. (2015)
		part/L	12.13	0.27	1.97	0.72	30.06	5724	1107.5	232.8	
		SA/L	−2.86	0.14	1.05	0.38	15.97	1567	230.5	76.9	
	Ag crustacean ($n = 31$)	g/L	−4.45	0.10	1.19	0.57	26.96	18,611	N/A	N/A	Blinova et al. (2013); Ribeiro et al. (2014); Silva et al. (2014); Ulm et al. (2015); Borase et al. (2019)
		part/L	12.54	0.19	2.19	1.05	49.67	113,077	507.6	84.2	
		SA/L	−2.78	0.12	1.42	0.68	32.15	34,025	82.8	19.2	
	TiO_2 algae ($n = 8$)	g/L	−1.81	0.18	1.32	0.51	76.44	9574	N/A	N/A	Ozkaleli and Erdem (2018)
		part/L	14.02	0.18	1.32	0.51	76.44	9574	0	0	
		SA/L	−0.53	0.18	1.32	0.51	76.44	9574	0	0	
	CuO crustacean ($n = 20$)	g/L	−1.62	0.20	1.95	0.90	26.19	6431	N/A	N/A	Heinlaan et al. (2008); Blinova et al. (2010); Manusadžianas et al. (2012); Kim et al. (2017); Rotini et al. (2018)
		part/L	14.29	0.26	2.49	1.16	33.52	23,856	270.9	28.0	
		SA/L	−0.13	0.23	2.16	1.00	29.04	9667	50.3	10.9	
Size Measured Inside of the Test Medium	ZnO bacteria ($n = 5$)	g/L	−1.07	0.04	0.29	0.08	13.55	281	N/A	N/A	Mallevre et al. (2014)
		part/L	12.72	0.04	0.29	0.08	13.55	281	0	0	
		SA/L	−0.29	0.04	0.29	0.08	13.55	281	0	0	
	Ag bacteria ($n = 18$)	g/L	−2.08	0.08	0.74	0.34	10.83	554	N/A	N/A	Mallevre et al. (2014); Mallevre et al. (2016)
		part/L	12.58	0.20	1.83	0.84	26.74	9054	1534.2	146.8	
		SA/L	−1.19	0.11	0.97	0.45	14.24	2010	262.9	31.4	
	Ag crustacean ($n = 24$)	g/L	−4.46	0.13	1.28	0.61	23.84	9115	N/A	N/A	Blinova et al. (2013); Silva et al. (2014); Ulm et al. (2015)
		part/L	10.53	0.15	1.51	0.71	28.08	36,069	295.7	17.8	
		SA/L	−3.46	0.10	0.97	0.46	17.99	5540	−39.2	−24.5	

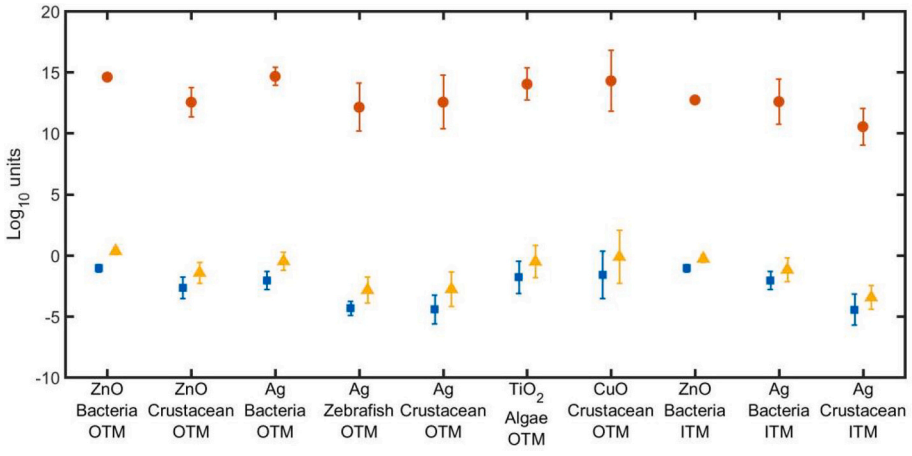

Fig. 3. Dispersion effects means and prediction interval for each subgroup in the dispersions effects analysis, blue squares represent the MPV data in units of log(g/L), yellow triangles represent SAPV units in log(m^2/L), and red circles represent the PPV units in log(particles/L). Error bars are present for all subgroups any missing are because they are too small to be seen. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

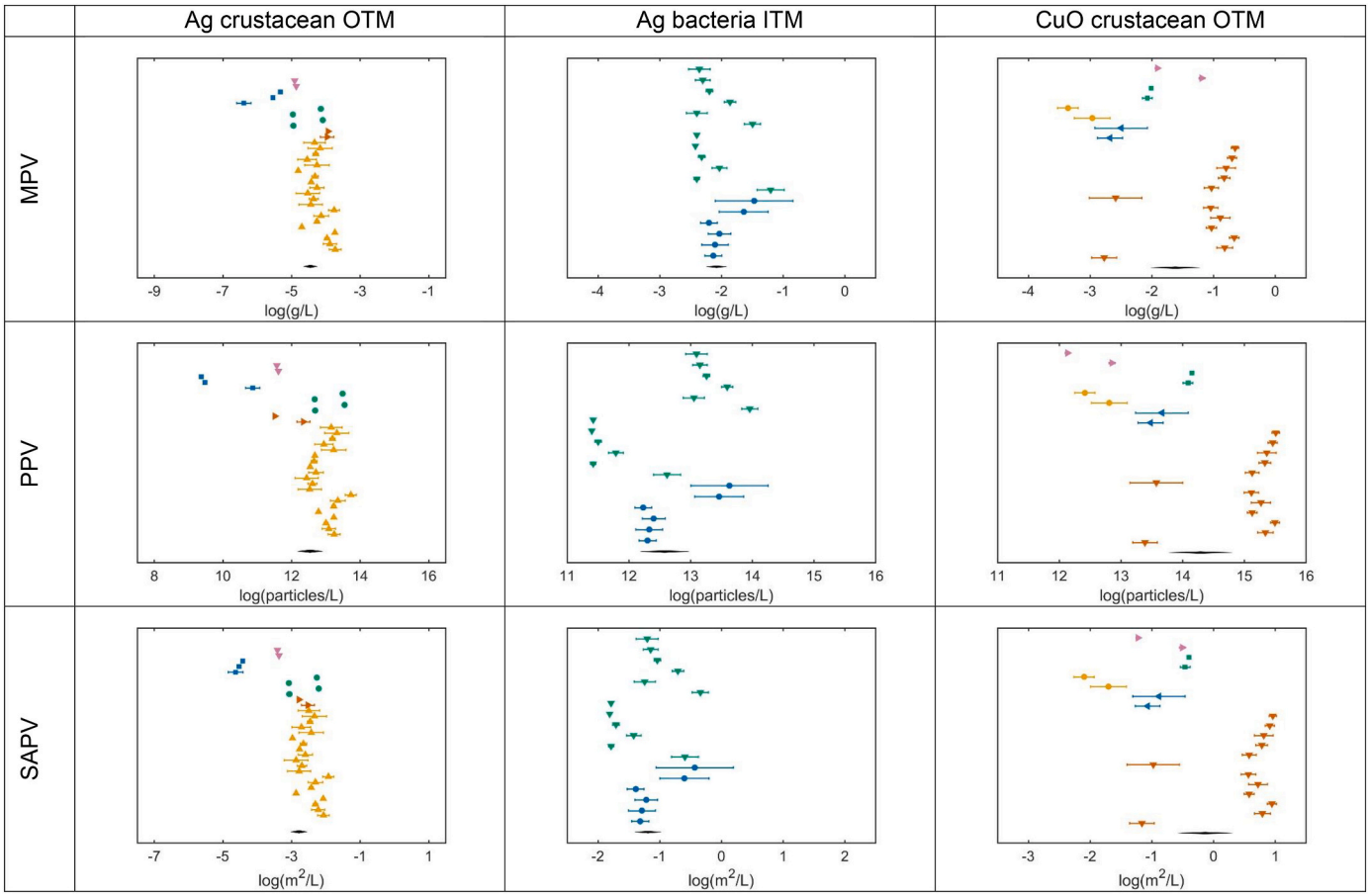

Fig. 4. Forest plots of the dispersion effects analysis for each combination of particle, organism, and chemistry as defined by the three units: grams per volume (MPV), particles per volume (PPV), and surface area per volume (SAPV). OTM = the size of the nanoparticle was measured out of the test medium. ITM = nanoparticle size determined in the test medium. Each point represents the mean and calculated 95% CI for each record. The black diamond (at bottom of each panel) is centered on the mean value from the meta-analysis where the width is the 95% CI values for the mean. The different colors correspond to the different sources of data. The Ag-crustacean-OTM column contains records from Blinova et al. (2013) represented by yellow upward-pointing triangles, Ribeiro et al. (2014) represented by green circles, Silva et al. (2014) represented by blue squares, Ulm et al. (2015) represented by a pink downward-pointing triangles, Borase et al. (2019) represented by orange right-pointing triangles. The Ag-bacteria-ITM column contains records from Mallevre et al. (2014) represented by a green downward-pointing triangle and Mallevre et al. (2016) represented by blue circles. The CuO-crustacean-OTM column contains records from Heinlaan et al. (2008) represented by a blue left-pointing triangle, Blinova et al. (2010) represented by orange downward-pointing triangles, Manusadžianas et al. (2012) represented by green squares, Kim et al. (2017) represented by yellow circles and Rotini et al. (2018) represented by pink right-pointing triangles. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

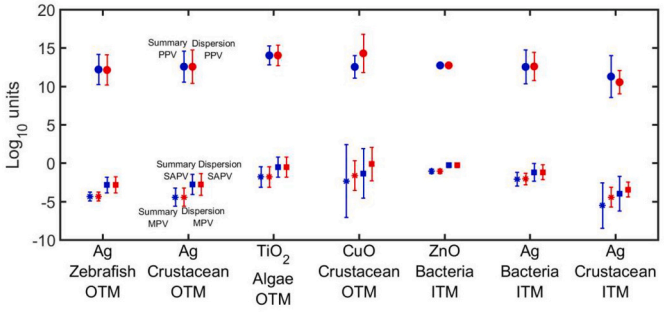

Fig. 5. Comparison of the summary effect (blue) and dispersion effect (red) analysis on the mean and 95% prediction intervals of each shared subgroup. Asterisks represent the MPV data in units of log(g/L), squares represent SAPV units in log(m^2/L), triangles represent the PPV units in log(particles/L). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

lack of change is attributable to the aforementioned lack of variation of size within the records in these two subgroups. In the case of the TiO_2/algae/OTM subgroup, all the records come from a single paper Ozkaleli and Erdem (2018), which varied the test medium and time but not the primary particle size. The CuO/crustacean/OTM did not show a preference for SAPV units as indicated by increases in Q (by 50.3%) and H (by 10.9%). However, this subgroup suffers from similar problems to the Ag/bacteria/OTM and TiO_2/algae/OTM, the nanoparticle size is too homogenous. Of the 20 records in these studies, 16 are reported as 30 nm, an additional 2 records are listed as 31.3 nm and the two remaining records are listed as 118 nm. As a result, the records are too similar for a clear signal to be seen. As MPV is converted to SAPV, the 118 nm particles studied by Rotini et al. (2018) (pink, right pointing triangle) move from a more central position to a more outlying position as seen in Fig. 4. This increases the heterogeneity because the remaining 30 nm particles remain fixed in their relative positions.

The Ag/bacteria subgroup are an interesting case study as the ITM and OTM versions contain the same 18 nanotoxicology tests. The two papers in these subgroups used the same particles, varying exposure time and the liquid media used in the toxicity tests. As a result of using the same particles, the Ag/bacteria/OTM subgroup showed no reduction

in heterogeneity for either the derived units, because the OTM size was the same for all records (i.e. 15 nm). Meanwhile the Ag/bacteria/ITM case shows that derived units increase heterogeneity in both SAPV and PPV. When looking at the forest plots of Fig. 4, we see that there are several clear outliers across the dataset, which correspond to toxicity reported at $t = 0$ and in 'crude wastewater' which skew the data, though even with the t = 0 records removed no reduction in heterogeneity is observed compared to MPV (Mallevre et al. 2014; Mallevre et al. 2016).

A similar comparison is possible for Ag/crustacean, which has sufficient records to compare across the various liquid media. Here, we find that the Ag/crustacean/ITM subgroup shows a 39.2 and 24.5% reduction in H and Q, respectively, for SAPV while the Ag/crustacean/OTM subgroup shows an 82.8 and 19.2% increase (Table 3). This striking dichotomy can be explained by looking at the forest plots for Ag/crustacean/OTM in Fig. 4. This shows data are well clustered for the MPV case. There is one strong outlier after units are converted to PPV and SAPV, while two additional records from the same study Silva et al. (2014) cluster together. If the three outlying records are removed, the remaining records show a 1.9% increase in H and 2.9% decrease in Q. This provides weak support for the hypothesis that SAPV is a better unit than MPV and evidence that the result may be data limited.

The outcome of the dispersion case analysis does not provide any support for or against the second part of the hypothesis, that is, non-dissolvable particles are best described by PPV units. This is because no data on non-soluble particles other than one TiO_2 subgroup met the inclusion criteria, though CeO_2 and other TiO_2 toxicity tests were also in the unfiltered dataset. Of the seven subgroups that are suitable for the dispersion analysis, two showed no change in PPV heterogeneity for reasons already described, and all of the other two subgroups showed an increase in the heterogeneity indicating that PPV is a less appropriate unit than MPV for generally dissolvable particles.

3.2.3. *Comparing ITM vs OTM*

Fig. 6 assesses the need for separating the data into subgroups based on the manner in which size is measured (i.e. measured *in* the test medium, ITM, or *out* of the test medium, OTM). This figure compares mean values for the dispersion effects analysis where each chemistry/organism pair have size data for both ITM and OTM. As expected, the MPV mean values are nearly identical because the studies within each subgroup contain overlapping if not exactly the same records. The derived units are different with PPV OTM having 2 log units (particles/L) more particles than the ITM data, and OTM SAPV having 0.65 log (m^2/L) more total surface area than ITM SAPV. These order of magnitude differences are substantial across the units and show that some consideration of where size is measured (i.e. in vs. out of test medium) is necessary. It is also worth noting that the prediction intervals for both methods overlap in several cases. This demonstrates the large range of concentration that is covered despite the fact that error was not propagated for the derived units in this analysis.

4. Discussion

There have been a number of recent calls (OECD (2017); Kraegeloh (2018)) to re-evaluate the ideal way to express the dose of nanoparticles in toxicology studies (e.g., grams per liter, surface area per liter, or number of particles per liter). Kennedy et al. (2015), for example, proposed that surface area is the best expression for dissolvable or photoreactive metallic particles such as Ag, ZnO, or TiO_2; whereas an expression of dose in terms of the number of particles would be best for insoluble particles such as CeO_2 and Au. They tested their hypothesis by exposing two organisms, a fish (*Pimephales promelas*) and a crustacean (*Ceriodaphnia dubia*), to Ag particles of various size. They determined that for both organisms the ion release, related to total surface area, provided the best fit in dose response curves.

The hypothesis of Kennedy et al. (2015) is supported by Warheit et al. (2007) who found that dose was better expressed in terms of surface reactivity for rat lungs dosed with different TiO_2 and SiO_2 nanoparticles despite both particles being insoluble. For most particles, surface reactivity is directly proportional to size but for TiO_2 the relationship is more fraught as rutile and anatase are often found in combination in larger nanoparticles and contribute differently to the particle's surface reactivity (Warheit 2010; Dobias and Bernier-Latmani 2013). Schmid and Stoeger (2016) confirmed that surface area is the optimal dose metric for aerosol nanomaterials for a wide variety of materials, recommending researchers prefer surface area over other dose metrics. Outside of metallic nanoparticles there is evidence that alternative units may be appropriate. Mottier et al. (2016) tested a variety of carbon nanoparticles on amphibians in water and found that surface area was the best expression for carbon nanoparticle doses, which was subsequently verified in a follow-up publication by Lagier et al. (2017).

While these past studies have shown evidence for alternative dosing metrics, each study has focused largely on *individual* tests. Here, we develop a new meta-analysis to help identify an optimal 'unit of choice' that is more broadly applicable to the toxicology of nanoparticles. This meta-analysis allows one to determine inter-study dispersion, account for the variation in data between studies, and provides a statistically rigorous yet simple framework for assessing the role of units in nanotoxicity.

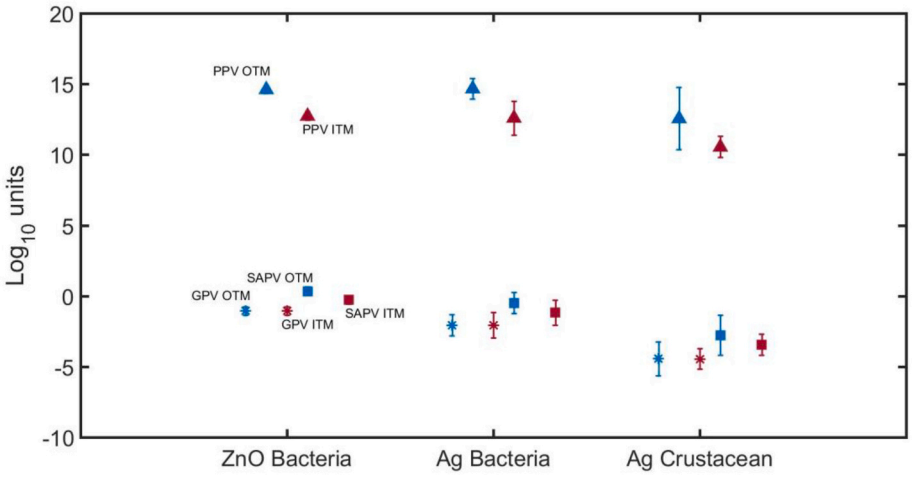

Fig. 6. Dispersion effects means and 95% prediction interval for each case that had paired ITM and OTM data. Blue are the OTM data points and magenta are the ITM data points. Asterisks represent the MPV data in units of log(g/L), squares represent SAPV units in log(m^2/L), and triangles represent PPV units in log (particles/L). Error bars are actually shown for ZnO: bacteria but they are too small to see. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.1. Understanding the effect of units

The overall results of our meta-analysis suggest that MPV (g/L) is *not* the best measure of toxicity and alternative units ought to be considered when analyzing the toxicity of nanoparticles. Specifically, our meta-analysis demonstrates that SAPV (m^2/L) is the better measure of mortality for soluble particles like Ag and ZnO for crustaceans. This was based on the reductions in heterogeneity compared to MPV (see Table 3). While SAPV was a better descriptor for Ag/crustacean/ITM and ZnO/crustacean/OTM, most other groups of dissolvable particles failed to provide sufficiently diverse data to provide a statistically meaningful conclusion with regard to SAPV. This shortcoming was *not* due to the meta-analysis. Rather, the vast majority of peer reviewed publications could not be utilized in the meta-analysis because they fail to report basic measures of statistics regarding particle size and concentration of putative nanotoxin (as discussed in more detail below).

Our meta-analysis was unable to test whether PPV is the best unit for non-soluble nanoparticles because much of the available data for particles like TiO_2 and CeO_2 failed to meet the inclusion requirements too (e.g., failed to report error in size measurements). We made no effort to exclude non-soluble particle chemistries from the dataset. The lack of a result for non-soluble particles indicates a clear knowledge gap in peer-reviewed publications and emphasizes the need for authors to thoroughly report as much metadata as possible.

The manner and medium (vacuum, air, liquid) in which particle size is measured are key caveats that should be addressed in future work. For example, size measurement performed by the experimentalist in the test medium (ITM) prior to adding the test organism differs from the nominal size reported by the manufacturer or the researcher's own TEM measurement (out of the test medium, OTM). Fig. 6 shows there is a difference in the mean derived units based on when and where size is determined. In our meta-analysis, we found no clear signal as to whether size measured ITM or OTM provide better results. From a perspective of reproducibility, OTM-size would be preferred because one would not need to make assumptions about size and shape when calculating the derived unit doses. However, it is unlikely that the OTM SAPV or PPV would be representative of actual exposure conditions. When exposed to environmental or physiological solutions, particles are known to aggregate and/or dissolve releasing ions that may interact with liquid components (e.g., culture medium) creating a plethora of different derivates among the original nanomaterial and its corona and therefore increasing the complexity of the system.

The real benefit of ITM-size is that it begins to account for particle aggregation. However, there are a few challenges for the use of ITM-size in the meta-analysis. For example, size would need to be determined as soon as particles were added to the test medium to minimize time for dissolution. Also, spherical shape would need to be assumed even in the case of aggregated particles. The effects of these assumptions are likely minimal on measures of heterogeneity, but they are expected to reduce utility of the calculated mean values. In either case, both ITM-size and OTM-size are imperfect measures of particle size because the size distribution will change over time and nanoparticle, being meta-stable, may tend towards their ionic form requiring more sophisticated methods to determine actual exposure (Bondarenko et al. 2013; Reidy et al. 2013; Mallevre et al. 2014; Kim et al. 2017). The time to develop better characterization methods for nanoparticles is now when predicted environmental concentrations of nanomaterials are orders of magnitude smaller than the mean toxic effects described here and have not yet been part of a largescale spill (Lazareva and Keller 2014; Garner et al. 2017).

4.2. Comparison of the new meta-analysis results to published review articles

Overall, the mean MPV values calculated in this meta-analysis compare favorably with reviews that summarize the existing nanotoxicology literature. The median estimates of Bondarenko et al. (2013) and Chen et al. (2015) shown in Fig. 7 and **Table S2** generally fall within the prediction intervals of our meta-analysis' estimate of mean toxic effect. Unfortunately, the prediction intervals cannot be compared to the spread of either study as their measured range was only reported pictorially.

There are a couple key pieces of insight to be gleaned from Fig. 7. First, the ZnO/bacteria case is one of two examples where there is no overlap between the meta-analytic mean, and the medians of Bondarenko et al. (2013) and Chen et al. (2015). Here, the meta-analytic mean was determined to be 0.76 log(g/L) lower than that of Bondarenko et al. (2013), but 0.55 log(g/L) higher than Chen et al. (2015). The differences between these mean values is likely due to the number and choice of records included in each of the estimates. The six records making up the ZnO/bacteria dataset in the meta-analysis are a subset of the 15 records of Bondarenko et al. (2013) dataset which itself is a subset of the 27 records from Chen et al. (2015). Moreover, both Chen et al. (2015) and the NanoE-Tox database (used herein) drew on the work of Bondarenko et al. (2013) to form the basis of their dataset. Therefore, it is not surprising that our meta-analysis estimates that the true toxic effect value falls between the two extremes (Juganson et al. 2015). The lack of agreement highlights a weakness of the present meta-analysis because a few records from limited numbers of studies biases the mean values even though the data quality standards may be quite stringent.

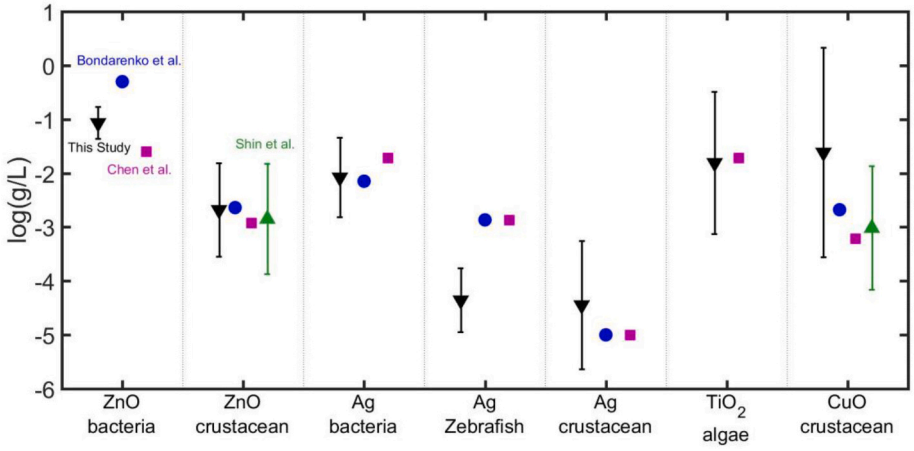

Fig. 7. Comparison of the g/L meta-analytic mean and prediction interval (black triangle) to the median values from Bondarenko et al. (2013) (blue circle) and Chen et al. (2015) (magenta square), and the mean and standard deviation values from Shin et al. (2018) (green triangle) for the subgroups in common with the meta-analysis. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 7 also illustrates the effect that "categorization" has on the dataset. Throughout this paper and published reviews, individual species are grouped together for the purpose of trying to make results more generalizable. However, which groupings are chosen can bias the data. For that reason, the largest appropriate subgroups were chosen for our meta-analysis. As an example of this approach, the Ag/Zebrafish subgroup in Fig. 7 is being compared to 'Ag/fish' subgroups in Bondarenko et al. (2013) and Chen et al. (2015), which partially explains the reason that their medians are not contained in the prediction interval. The choice was made to use the more specific name classifier because all data that met inclusion criteria were actually zebrafish and both subgroups were being used interchangeably in the NanoE-Tox database (Juganson et al. 2015). For another example, the NanoE-Tox database was found to use both crustacean and the more specific copepod category, which were combined into one crustacean category in our meta-analysis.

Fig. 7 and **Table S2** also present results from Shin et al. (2018), another review that also performed a meta-analysis, although it focused only on the crustacean species *Daphnia magna*, a common test organism. Their work produced several mean values for many different types of nanoparticles along with the calculated standard deviation, which is presented as error bars in Fig. 7. It is unclear how the studies were weighted, if at all. Once again, Shin et al. (2018) show good agreement with the work presented in this meta-analysis. Their standard deviation is similar to the prediction intervals presented in our meta-analysis. It is also worth highlighting that Shin et al. (2018) calculated mean Ag nanoparticle toxicity but broke those data points out by coating. As such, their mean and standard deviation for Ag are not presented in Fig. 7. However, their mean values range from −4.02 log(g/L) to −2.41 log(g/L), which overlaps with our prediction interval of the Ag/crustacean though their estimates extend to much higher concentrations.

4.3. Limitations and directions of future work

The meta-analysis described in this paper provides a statistically rigorous yet simple approach to determine which units are best suited for expressing toxicity specifically for nanoparticles. An important limitation, and a critical point to highlight, is the lack of reliable data in peer-reviewed publications that can be input to this meta-analysis. The effect of units on mortality endpoints could be rigorously determined for only two subgroups (Ag/crustacean/ITM and ZnO/crustacean/OTM) even when using the largest, publicly available database with greater than 1500 records on nanoparticle toxicity, the NanoE-Tox database, combined with an up-to-date literature search. It is also important to highlight that, at present, the NanoE-Tox database is the only publicly available database of nanotoxicology data. When we looked for other databases we found that most are nothing more than a literature matrix, the database itself is no longer updated or it is lost behind dead links (e.g., Nanomaterial-Biological Interations Knowledgebase (2010); International Council on Nanotechnology (2014); Chen et al. (2015); Maimon and Browarnik (2018)).

The conclusions possible with our meta-analysis are *not* limited by the quantity of data but the lack of standardization of data. This has less to do with the authors of databases like NanoE-Tox and more to do with the primary literature's failure to report basic values like uncertainty in measurements of size and toxicity. The biggest limit on data inclusion, and *a point where great strides can be taken*, is ensuring that meaningful error descriptors are reported for (i) EC_{50} concentration and (ii) nanoparticle size.

One of the subtleties that was found as part of the meta-analysis was that many of the studies in the database reported the 95% confidence interval (95% CI) of EC_{50} but not the standard deviation, which is necessary because meta-analysis weights studies based on inverse variance. Complicating matters, there are two methods used in the literature to calculate the 95% CI. The first is the simple t-statistic method taught in introductory statistics classes and produces a symmetrical interval; whereas the second method uses a bootstrapping approach recommended by the EPA, which calculates a nonsymmetrical 95% CI (Norberg-King 1993). In either case, standard deviation cannot be recovered, and otherwise quality studies had to be excluded from our analysis because error was not reported in an interoperable manner. Therefore, we recommend researchers report standard deviation along with the 95% confidence interval.

We also found variability in the error descriptor for particle size, with particle size being reported in four different ways in the NanoE-Tox database. As an example, a "15 nm" particle could be reported as: 15 nm, 15 ± 5 nm, 10–20 nm, or < 25 nm. Only one of these, 15 ± 5 nm, is sufficient for inclusion in the summary effects analysis while the dispersion case also allowed the 15 nm nominal size. The 15 ± 5 format is preferred because it offers both the actual size of the particle synthesized and the dispersion around the mean. This is important because no current synthesis method for metallic nanoparticle creates perfectly repeatable or monodisperse product and minute size differences can change the results of the analysis (Liu et al. 2010).

Reporting size as <25 nm simply means a particle is in fact a nanoparticle. It provides no estimate of the error in size. Reporting size as 10–20 nm is deceiving because it can and has been assumed to be equivalent to 15 ± 5 nm, but it is not (Hua et al. 2016). After contacting several manufacturers who report size as 10–20 nm, we learned that this form indicates the mean particle size within the range and no dispersion about the mean can be assumed. It should be noted the manufacturers of these particles tend to be industrial producers, and the particles are not intended for scientific research. If these industrial sources are going to be used for nanoparticle research, the particles should have their actual size distribution quantified and reported as strongly advised in various guidelines (Thomas et al. 2013; Oksel et al. 2015; Marchese Robinson et al. 2016; Chen et al. 2017; OECD 2017).

Others have written extensively on what data should be included outside of the measured toxic effect and their recommendations are echoed here. First, as Marchese Robinson et al. (2016) recognized, every researcher has their own theories about what properties are important and satisfying everyone is not feasible because as we gain more knowledge the "essential" data points will change. However, what is unlikely to change are the four agreed upon minimum nanoparticle properties that should be measured and reported in any nanotoxicology study: composition, shape, crystallinity, and initial size at manufacturer (Marchese Robinson et al. 2016). Based on our meta-analysis, we would modify this list by adding a fifth property, the error in initial size. We also encourage independent verification of the size of the particle (e.g., confirm size and error that is reported by the manufacturer).

Based on the results presented herein, we recommend that future studies at least examine the effect of units (MPV vs. SAPV vs. PPV) on their own data, something that does not even require a meta-analysis. Future efforts could also use our meta-analytic framework to look at endpoints beyond mortality and place a greater emphasis on meta-regression to re-examine properties such as the coating on a particle, length of exposure, and expand to assess the effect of shape of the particle, the method used to size the particle, and biological test species as these attributes are generally reported in publications. For example, one could perform meta-regression to assess the effect of test species, building on the study of Shin et al. (2018) on *Daphnia magna*, or determine how shape, e.g., nano-needles vs. nano-spheres impacts toxicity building on Hua et al. (2016). Further, we recommend (i) researchers measure and report the error in nanoparticle size in the main text body and (ii) report error in toxicity measurement as both 95% confidence interval and as standard deviation. Following these recommendations will allow for greater data interoperability as the field advances.

Funding sources

We acknowledge support from the U.S. National Science Foundation grant EAR-2038207 and the Friends of Orton Hall fund at Ohio State University.

Declaration of Competing Interest

None.

Acknowledgements

We acknowledge the constructive criticism of five anonymous reviewers and Sarah Wheeler for proofreading the article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.impact.2020.100277.

References

International Council on Nanotechnology, 2014. Nano-EHS Database Analysis Tool.
Nanomaterial-Biological Interations Knowledgebase, 2010. NBI Knowledgebase.
Alkilany, A.M., Mahmoud, N.N., Hashemi, F., Hajipour, M.J., Farvadi, F., Mahmoudi, M., 2016. Misinterpretation in Nanotoxicology: a personal perspective. Chem. Res. Toxicol. 29 (6), 943–948. https://doi.org/10.1021/acs.chemrestox.6b00108.
Arvidsson, R., 2018. Risk assessments show engineered nanomaterials to be of low environmental concern. Environ. Sci. Technol. 52 (5), 2436–2437. https://doi.org/10.1021/acs.est.8b00754.
Blanco, E., Shen, H., Ferrari, M., 2015. Principles of nanoparticle design for overcoming biological barriers to drug delivery. Nat. Biotechnol. 33 (9), 941–951. https://doi.org/10.1038/nbt.3330.
Blinova, I., Ivask, A., Heinlaan, M., Mortimer, M., Kahru, A., 2010. Ecotoxicity of nanoparticles of CuO and ZnO in natural water. Environ. Pollut. 158 (1), 41–47. https://doi.org/10.1016/j.envpol.2009.08.017.
Blinova, I., Niskanen, J., Kajankari, P., Kanarbik, L., Käkinen, A., Tenhu, H., Penttinen, O.-P., Kahru, A., 2013. Toxicity of two types of silver nanoparticles to aquatic crustaceans Daphnia magna and Thamnocephalus platyurus. Environ. Sci. Pollut. Res. 20 (5), 3456–3463. https://doi.org/10.1007/s11356-012-1290-5. http://link.springer.com/10.1007/s11356-012-1290-5. (Accessed 29 June 2016).
Boehme, S., Staerk, H.-J., Reemtsma, T., Kuehnel, D., 2015. Effect propagation after silver nanoparticle exposure in zebrafish (Danio rerio) embryos: a correlation to internal concentration and distribution patterns. Environ. Sci. 2 (6), 603–614. https://doi.org/10.1039/c5en00118h.
Bondarenko, O., Juganson, K., Ivask, A., Kasemets, K., Mortimer, M., Kahru, A., 2013. Toxicity of Ag, CuO and ZnO nanoparticles to selected environmentally relevant test organisms and mammalian cells in vitro: a critical review. Arch. Toxicol. 87 (7), 1181–1200. https://doi.org/10.1007/s00204-013-1079-4.
Borase, H.P., Patil, S.V., Singhal, R.S., 2019. Moina macrocopa as a non-target aquatic organism for assessment of ecotoxicity of silver nanoparticles: effect of size. Chemosphere. 219, 713–723. https://doi.org/10.1016/j.chemosphere.2018.12.031.
Bundschuh, M., Seitz, F., Rosenfeldt, R.R., Schulz, R., 2016. Effects of nanoparticles in fresh waters: risks, mechanisms and interactions. Freshw. Biol. 61 (12), 2185–2196. https://doi.org/10.1111/fwb.12701.
Bundschuh, M., Filser, J., Lüderwald, S., McKee, M.S., Metreveli, G., Schaumann, G.E., Schulz, R., Wagner, S., 2018. Nanoparticles in the environment: where do we come from, where do we go to? Environ. Sci. Eur. 30 (1), 6. https://doi.org/10.1186/s12302-018-0132-6.
Caballero-Guzman, A., Nowack, B., 2016. A critical review of engineered nanomaterial release data: are current data useful for material flow modeling? Environ. Pollut. 213, 502–517. https://doi.org/10.1016/j.envpol.2016.02.028.
Chen, G., Vijver, M.G., Peijnenburg, W.J.G.M., 2015. Summary and analysis of the currently existing literature data on metal-based nanoparticles published for selected aquatic organisms: applicability for toxicity prediction by (Q)SARs. Altern. Lab. Anim 43 (4), 221–240. https://doi.org/10.1177/026119291504300404.
Chen, G., Roy, I., Yang, C., Prasad, P.N., 2016. Nanochemistry and nanomedicine for nanoparticle-based diagnostics and therapy. Chem. Rev. 116 (5), 2826–2885. https://doi.org/10.1021/acs.chemrev.5b00148.
Chen, G., Peijnenburg, W., Xiao, Y., Vijver, M., 2017. Current knowledge on the use of computational toxicology in hazard assessment of metallic engineered nanomaterials. Int. J. Mol. Sci. 18 (7), 1504. https://doi.org/10.3390/ijms18071504.
Dobias, J., Bernier-Latmani, R., 2013. Silver release from silver nanoparticles in natural waters. Environ. Sci. Technol. 47 (9), 4140–4146. https://doi.org/10.1021/es304023p.
Elsaesser, A., Howard, C.V., 2012. Toxicology of nanoparticles. Adv. Drug Deliv. Rev. 64 (2), 129–137. https://doi.org/10.1016/j.addr.2011.09.001.
Garner, K.L., Suh, S., Keller, A.A., 2017. Assessing the risk of engineered nanomaterials in the environment: development and application of the nanoFate model. Environ. Sci. Technol. 51 (10), 5541–5551. https://doi.org/10.1021/acs.est.6b05279.
Heinlaan, M., Ivask, A., Blinova, I., Dubourguier, H.-C., Kahru, A., 2008. Toxicity of nanosized and bulk ZnO, CuO and TiO2 to bacteria Vibrio fischeri and crustaceans Daphnia magna and Thamnocephalus platyurus. Chemosphere. 71 (7), 1308–1316. https://doi.org/10.1016/j.chemosphere.2007.11.047.
Higgins, J.P.T., Thompson, S.G., 2002. Quantifying heterogeneity in a meta-analysis. Stat. Med. 21 (11), 1539–1558. https://doi.org/10.1002/sim.1186.
Hua, J., Vijver, M.G., Chen, G., Richardson, M.K., Peijnenburg, W.J.G.M., 2016. Dose metrics assessment for differently shaped and sized metal-based nanoparticles. Environ. Toxicol. Chem. 35 (10), 2466–2473. https://doi.org/10.1002/etc.3414.
Jemec, A., Kahru, A., Potthoff, A., Drobne, D., Heinlaan, M., Böhme, S., Geppert, M., Novak, S., Schirmer, K., Rekulapally, R., et al., 2016. An interlaboratory comparison of nanosilver characterisation and hazard identification: Harmonising techniques for high quality data. Environ. Int. 87 (July 2015), 20–32. https://doi.org/10.1016/j.envint.2015.10.014.
Juganson, K., Ivask, A., Blinova, I., Mortimer, M., Kahru, A., 2015. NanoE-Tox: new and in-depth database concerning ecotoxicity of nanomaterials. Beilstein J. Nanotechnol. 6 (1), 1788–1804. https://doi.org/10.3762/bjnano.6.183.
von der Kammer, F., Ferguson, P.L., Holden, P.A., Masion, A., Rogers, K.R., Klaine, S.J., Koelmans, A.A., Horne, N., Unrine, J.M., 2012. Analysis of engineered nanomaterials in complex matrices (environment and biota): general considerations and conceptual case studies. Environ. Toxicol. Chem. 31 (1), 32–49. https://doi.org/10.1002/etc.723.
Keller, A.A., Adeleye, A.S., Conway, J.R., Garner, K.L., Zhao, L., Cherr, G.N., Hong, J., Gardea-Torresdey, J.L., Godwin, H.A., Hanna, S., et al., 2017. Comparative environmental fate and toxicity of copper nanomaterials. NanoImpact. 7 (May), 28–40. https://doi.org/10.1016/j.impact.2017.05.003.
Kennedy, A.J., Hull, M.S., Diamond, S., Chappell, M., Bednar, A.J., Laird, J.G., Melby, N. L., Steevens, J.A., 2015. Gaining a critical mass: a dose metric conversion case study using silver nanoparticles. Environ. Sci. Technol. 49 (20), 12490–12499. https://doi.org/10.1021/acs.est.5b03291.
Kim, S., Samanta, P., Yoo, J., Kim, W.-K., Jung, J., 2017. Time-dependent toxicity responses in Daphnia magna exposed to CuO and ZnO nanoparticles. Bull. Environ. Contam. Toxicol. 98 (4), 502–507. https://doi.org/10.1007/s00128-016-2022-1.
Klaine, S.J., Alvarez, P.J.J., Batley, G.E., Fernandes, T.F., Handy, R.D., Lyon, D.Y., Mahendra, S., McLaughlin, M.J., Lead, J.R., 2008. Nanomaterials in the environment: behavior, fate, bioavailability, and effects. Environ. Toxicol. Chem. 27 (9), 1825–1851. https://doi.org/10.1897/08-090.1.
Kraegeloh, A., 2018. Nanosafety research–an ongoing story. Chem. Res. Toxicol. 31 (11), 1105. https://doi.org/10.1021/acs.chemrestox.8b00325.
Lagier, L., Mouchet, F., Laplanche, C., Mottier, A., Cadarsi, S., Evariste, L., Sarrieu, C., Lonchambon, P., Pinelli, E., Flahaut, E., et al., 2017. Surface area of carbon-based nanoparticles prevails on dispersion for growth inhibition in amphibians. Carbon N Y. 119, 72–81. https://doi.org/10.1016/j.carbon.2017.04.016.
Lazareva, A., Keller, A.A., 2014. Estimating potential life cycle releases of engineered nanomaterials from wastewater treatment plants. ACS Sustain. Chem. Eng. 2 (7), 1656–1665. https://doi.org/10.1021/sc500121w.
Liu, W., Wu, Y., Wang, C., Li, H.C., Wang, T., Liao, C.Y., Cui, L., Zhou, Q.F., Yan, B., Jiang, G.B., 2010. Impact of silver nanoparticles on human cells: effect of particle size. Nanotoxicology. 4 (3), 319–330. https://doi.org/10.3109/17435390.2010.483745.
Maimon, O., Browarnik, A., 2018. NHECD.
Mallevre, F., Fernandes, T.F., Aspray, T.J., 2014. Silver, zinc oxide and titanium dioxide nanoparticle ecotoxicity to bioluminescent Pseudomonas putida in laboratory medium and artificial wastewater. Environ. Pollut. 195, 218–225. https://doi.org/10.1016/j.envpol.2014.09.002.
Mallevre, F., Alba, C., Milne, C., Gillespie, S., Fernandes, T.F., Aspray, T.J., 2016. Toxicity testing of pristine and aged silver nanoparticles in real wastewaters using bioluminescent Pseudomonas putida. Nanomaterials. 6 (3), 49. https://doi.org/10.3390/nano6030049.
Manusadžianas, L., Caillet, C., Fachetti, L., Gylytė, B., Grigutytė, R., Jurkonienė, S., Karitonas, R., Sadauskas, K., Thomas, F., Vitkus, R., et al., 2012. Toxicity of copper oxide nanoparticle suspensions to aquatic biota. Environ. Toxicol. Chem. 31 (1), 108–114. https://doi.org/10.1002/etc.715.
Marchese Robinson, R.L., Lynch, I., Peijnenburg, W., Rumble, J., Klaessig, F., Marquardt, C., Rauscher, H., Puzyn, T., Purian, R., Åberg, C., et al., 2016. How should the completeness and quality of curated nanomaterial data be evaluated? Nanoscale. 8 (19), 9919–9943. https://doi.org/10.1039/C5NR08944A.
Mottier, A., Mouchet, F., Laplanche, C., Cadarsi, S., Lagier, L., Arnault, J.-C., Girard, H. A., León, V., Vázquez, E., Sarrieu, C., et al., 2016. Surface area of carbon nanoparticles: a dose metric for a more realistic ecotoxicological assessment. Nano Lett. 16 (6), 3514–3518. https://doi.org/10.1021/acs.nanolett.6b00348.
Norberg-King, T., 1993. A linear interpolation method for sublethal toxicity: the inhibition concentration (ICp) approach. Version 2.0. Natl. Effl. Toxic. Assess. Cent. Tech. Rep. 39 (July).
OECD, 2017. Draft Guidance Document on Aquatic and Sediment Toxicological Testing of Nanomaterials (Organisation).
Oksel, C., Ma, C.Y., Wang, X.Z., 2015. Current situation on the availability of nanostructure–biological activity data. SAR QSAR Environ. Res. 26 (2), 79–94. https://doi.org/10.1080/1062936X.2014.993702.
Ozkaleli, M., Erdem, A., 2018. Biotoxicity of TiO2 nanoparticles on raphidocelis subcapitata microalgae exemplified by membrane deformation. Int. J. Environ. Res. Public Health 15 (3), 416. https://doi.org/10.3390/ijerph15030416.
Qu, X., Alvarez, P.J.J., Li, Q., 2013. Applications of nanotechnology in water and wastewater treatment. Water Res. 47 (12), 3931–3946. https://doi.org/10.1016/j.watres.2012.09.058.
Reidy, B., Haase, A., Luch, A., Dawson, K., Lynch, I., 2013. Mechanisms of silver nanoparticle release, transformation and toxicity: a critical review of current knowledge and recommendations for future studies and applications. Materials (Basel). 6 (6), 2295–2350. https://doi.org/10.3390/ma6062295.
Ribeiro, F., Gallego-Urrea, J.A., Jurkschat, K., Crossley, A., Hassellöv, M., Taylor, C., Soares, A.M.V.M., Loureiro, S., 2014. Silver nanoparticles and silver nitrate induce high toxicity to Pseudokirchneriella subcapitata, *Daphnia magna* and *Danio rerio*. Sci.

Total Environ. 466–467, 232–241. https://doi.org/10.1016/j.scitotenv.2013.06.101.

Rotini, A., Gallo, A., Parlapiano, I., Berducci, M.T., Boni, R., Tosti, E., Prato, E., Maggi, C., Cicero, A.M., Migliore, L., et al., 2018. Insights into the CuO nanoparticle ecotoxicity with suitable marine model species. Ecotoxicol. Environ. Saf. 147, 852–860. https://doi.org/10.1016/j.ecoenv.2017.09.053.

Salieri, B., Turner, D.A., Nowack, B., Hischier, R., 2018. Life cycle assessment of manufactured nanomaterials: where are we? NanoImpact. 10 (November 2017), 108–120. https://doi.org/10.1016/j.impact.2017.12.003.

Saunders, B.R., Turner, M.L., 2008. Nanoparticle–polymer photovoltaic cells. Adv. Colloid Interf. Sci. 138 (1), 1–23. https://doi.org/10.1016/j.cis.2007.09.001.

Schmid, O., Stoeger, T., 2016. Surface area is the biologically most effective dose metric for acute nanoparticle toxicity in the lung. J. Aerosol Sci. 99, 133–143. https://doi.org/10.1016/j.jaerosci.2015.12.006.

Schrurs, F., Lison, D., 2012. Focusing the research efforts. Nat. Nanotechnol. 7 (9), 546–548. https://doi.org/10.1038/nnano.2012.148.

Shin, H.K., Seo, M., Shin, S.E., Kim, K., Park, J., No, K.T., 2018. Meta-analysis of Daphnia magna nanotoxicity experiments in accordance with test guidelines. Environ. Sci. Nano. 5 (3), 765–775. https://doi.org/10.1039/C7EN01127J.

Silva, T., Pokhrel, L.R., Dubey, B., Tolaymat, T.M., Maier, K.J., Liu, X., 2014. Particle size, surface charge and concentration dependent ecotoxicity of three organo-coated silver nanoparticles: comparison between general linear model-predicted and observed toxicity. Sci. Total Environ. 468–469, 968–976. https://doi.org/10.1016/j.scitotenv.2013.09.006.

Steinhäuser, K.G., Sayre, P.G., 2017. Reliability of methods and data for regulatory assessment of nanomaterial risks. NanoImpact. 7 (April), 66–74. https://doi.org/10.1016/j.impact.2017.06.001.

Thomas, D.G., Gaheen, S., Harper, S.L., Fritts, M., Klaessig, F., Hahn-Dantona, E., Paik, D., Pan, S., Stafford, G.A., Freund, E.T., et al., 2013. ISA-TAB-nano: a specification for sharing nanomaterial research data in spreadsheet-based format. BMC Biotechnol. 13 (1), 2. https://doi.org/10.1186/1472-6750-13-2.

Ulm, L., Krivohlavek, A., Jurasin, D., Ljubojevic, M., Sinko, G., Crnkovic, T., Zuntar, I., Sikic, S., Vrcek, I.V., 2015. Response of biochemical biomarkers in the aquatic crustacean Daphnia magna exposed to silver nanoparticles. Environ. Sci. Pollut. Res. 22 (24), 19990–19999. https://doi.org/10.1007/s11356-015-5201-4.

Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. J. Stat. Softw. 36 (3), 1–48 doi:10.18637/jss.v036.i03.

Warheit, D.B., 2010. Debunking some misconceptions about nanotoxicology. Nano Lett. 10 (12), 4777–4782. https://doi.org/10.1021/nl103432w.

Warheit, D.B., Webb, T.R., Reed, K.L., Frerichs, S., Sayes, C.M., 2007. Pulmonary toxicity study in rats with three forms of ultrafine-TiO2 particles: differential responses related to surface properties. Toxicology. 230 (1), 90–104. https://doi.org/10.1016/j.tox.2006.11.002.