Supporting Information

ADAPTING OECD AQUATIC TOXICITY TESTS FOR USE WITH MANUFACTURED NANOMATERIALS: KEY ISSUES AND CONSENSUS RECOMMENDATIONS

Elijah J. Petersen[1], Stephen A. Diamond[2], Alan J. Kennedy[3,*], Greg G. Goss[4], Kay Ho[5], Jamie Lead[6], Shannon K. Hanna[1], Nanna B. Hartmann[7], Kerstin Hund-Rinke[8], Brian Mader[9], Nicolas Manier[10], Pascal Pandard[10], Edward R. Salinas[11], Phil Sayre[12]

*** Corresponding author:** Alan.J.Kennedy@usace.army.mil; phone: 601-634-3344; fax: 601-634-2263

[1] Biosystems and Biomaterials Division, Material Measurement Laboratory, National Institute of Standards and Technology, Gaithersburg, MD, USA
[2] NanoSafe, Inc, Midwest, Duluth, MN, USA
[3] US Army Engineer Research and Development Center, Environmental Laboratory, Vicksburg, MS, USA
[4] Department of Biological Sciences, University of Alberta and National Institute of Nanotechnology, National Research Council, Edmonton, Alberta, Canada
[5] Office of Research and Development, National Health and Environmental Effects Research Laboratory – Atlantic Ecology Division, United States Environmental Protection Agency, Narragansett, RI, USA
[6] Center for Environmental Nanoscience and Risk, Department of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC 29036, USA
[7] Department of Environmental Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark
[8] Fraunhofer Institute for Molecular Biology and Applied Ecology, Schmallenberg, Germany
[9] 3M, Environmental Laboratoary, St. Paul, MN, USA
[10] Institute National de l'Environnement Industriel et des Risques (INERIS), Parc Technologique ALATA, F 60 550 Verneuil en-Halatte, France.
[11] Experimental Toxicology and Ecology, BASF SE, Ludwigshafen, Germany
[12] formerly with Office of Pollution Prevention and Toxics, United States Environmental Protection Agency, Washington D.C., USA., phil.sayre@verizon.net

6 pages, 1 table

Supplemental discussion of definitions and measurements of “dissolved” substances.

The OECD GD on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media recommends a 0.2 µm filter to determine the dissolved fraction,[1] while the American Public Health Association recommends using a 0.45 µm pore-size filter.[2-4] More recently, size discrimination at 1 nm and 1 mm has been used in agreement with colloid chemistry and environmentally relevant processes. However in practice, this upper limit is often reduced to 0.1 mm to 0.2 mm to exclude biological material. In addition, there are processes, such as the impacts of pH, flow rate, pressure, clogging of pores, and losses from sorption of dissolved species to the membranes, which complicate interpretation of data from these methods. From a molecular point of view, and independent of operational definitions, a substance is defined as dissolved when individual molecules of the solute are separated by the molecules of a solvent[5] whereby solid-state MNs in dispersion clearly do not fall under this definition.

Supplemental discussion regarding which type of natural organic matter (NOM) to recommend for aquatic toxicity testing

Advantages of commercially available humic acid are that large masses can be purchased and it is less expensive than the more standardized materials (e.g., Suwanee River Fulvic Acid (SRFA) available from the International Humic Substances Society). However, commercially-available NOM is generally not extracted using transparent and standardized methods and sources, generally have high ash contents, are more heterogeneous, and may be isolated from soils, unlike SRFA. In addition, NOM from aquatic sources does vary substantially,[6, 7] and thus SRFA, a single NOM type, does not reflect the range of NOM characteristics in natural waters. Overall, a consensus was not reached about which type of NOM to recommend.

Supplemental discussion of the impact of calculating toxicity endpoints if the 20% specification is not achieved

Generally OECD guidance advocates using the geometric (for static and semi-static exposures) or arithmetic (for flow through exposures) mean measured concentration when there are deviations from the ± 20 % exposure specification.[8] Two TGs, OECD 201 Freshwater Alga and Cyanobacteria, Growth Inhibition Test,[9] and OECD 211 *Daphnia magna* Reproduction Test,[10] offer specific guidance on calculating the toxicity endpoint if the aqueous phase concentration changes by more than 20 % of the initial concentration. Both of these TGs describe calculation of exposure using the geometric mean of measured values when the exposure variation exceeds ± 20 %. OECD 211[10] illustrates the time weighted average approach using an example (Annex 6) where the concentration of a tested substance decreases and increases between water renewals. Exposure estimates are derived from the following equation to determine the time-weighted-average:

$$Area = \frac{[C_0 - C_1]}{[Ln(C_0) - Ln(C_1)]} \bullet D$$

> *Where C_0 = measured concentration immediately after chemical addition, C_1 = measured concentration at the end of the renewal period, D = the number of renewal period days*

This approach assumes exponential losses, and it is noted that other types of kinetic losses may require alternate models. The potential errors in exposure-response estimation that arise if material losses are greater than assumed, or if loss kinetics differ from assumed models (e.g. linear or logarithmic models) have been previously quantitatively discussed.[9, 11] The various scenarios depicted are for algal, 72 h bioassays where exposure-response models are based on experiments where toxicant losses were manipulated. Inhibitory concentrations (ICs) of 15 %, 25 %, and 50 % were calculated for 24 h, 48 h, and 72 h exposures for nine substance-loss scenarios including no-loss (the accurate exposure-response estimate used for comparison), linear, slightly curved, and exponential losses. Examples are provided for cases in which the concentrations fall to zero well before the end of the test. IC estimates for each loss scenario can then be compared to the no-loss scenario; representing the outcomes when losses are not accounted for. When modeled losses are about 40 % and unaccounted for, IC50 estimates for 72 h exposures are 1.4x and 1.7x higher for linear and curved kinetics, respectively. When concentrations fall to nearly zero at 72 h, IC_{50} estimates are 2.2x, 3.0x, and 6.0x higher for linear, curved, and exponential losses, respectively. When losses are exponential and reach 95 % by 48 h or 36 h, IC_{50} estimates are 15.7x and 42.8x higher, respectively. In the scenario where the losses are exponential and decrease to 95 % after 36 h, using an average concentration based on initial and final values and the assumption of linear loss kinetics reduces the IC_{50} estimate error to 25x. This approach is an improvement, but also illustrates the need for concentration measurements at sufficiently frequent intervals to provide for more accurate exposure estimation. Finally, it is important to recognize that the 20 % exposure specification is based wholly on substance-mass concentrations. While this metric is certainly applicable to MNs, other metrics such as particle number concentration, size distribution, or surface area concentration may prove to be equally, or more, important with regards to accurately assessing MN ecological effects. Important questions are whether exposure thresholds for these MN-specific metrics also need to be evaluated and how to incorporate endpoints based on these metrics into environmental risk assessment frameworks.[12]

Table S1. Key additional research topics to fulfill data gaps needed for prescriptive use of OECD aquatic toxicity methods for nanomaterial testing organized by manuscript section

Section	Key research topics
Key considerations related to NM aquatic toxicity testing	• Research could support development of a single test medium for each TG that would lead to the most reliable ecotoxicity results for MN testing. One important consideration when choosing the medium is to avoid nutrient depletion from high concentrations of MNs with high sorption coefficients (see [13] for additional details), which would influence the bioassay results. However, there may be limitations as discussed above with regards to changing the test media such as the potential impact on organism health and decreased environmental relevance that must be considered. In all cases, the validity criteria of the selected method must be fulfilled in the chosen test medium. • Different types of test containers (size, type of material, geometry) should be tested to assess the robustness of the different TGs with regard to this parameter. Significant sorption onto the sidewalls of test containers could lead to an underestimation of toxicity. Also, the impact of the stirring protocol should be evaluated for tests, such as required by the algae growth inhibition test,[9] since stirring will likely impact MN agglomeration and potentially settling. • Assess the within and between laboratory reliability of testing the toxicity of the initial suspension vs. pseudo steady state suspension with a MN present. These data are necessary to inform more prescriptive guidance regarding how to best prepare the MN dispersion in the test media for reliable testing. • Identification and minimization of potential artifacts when using OECD TGs to test MNs is a key topic. This information should be assessed for each TG and the most reliable analytical method, such as testing algal concentrations using fluorometric measurements,[14] needs to be provided.
Considerations for applying the 20 % exposure specification to testing NMs	• As with all other chemicals, the possibility of conforming to the 20% exposure specification should be investigated prior to undertaking MN bioassays. Research should focus on approaches to maintaining MNs in suspension, including (but not limited to) more frequent renewal of media, use of flow-through delivery systems, media composition and pH and addition of NOM. • Investigate methods of maintaining stable exposure concentrations (± 20%) and models to determine appropriate exposure concentrations under changing conditions. Testing of flow through systems should include investigation of the potential for increased MN concentrations in the test system resulting from settled material not removed from chambers in out-flowing media. A related research need is the development of reliable, rapid and preferably standardized methods to quantify MNs and their transformation in aquatic test media throughout the duration of the tests. • Research is needed to determine if maintaining stable concentrations reduces variability in test results when agglomeration and dissolution cannot be avoided. • Research is needed to determine the extent to which settled MNs influence ecotoxicity results. If settled MNs have a significant impact on water-column test results, it might be possible to isolate settled material from test organisms using screens or other apparatus to prevent organisms from interacting with the settled MNs. Corollary questions are whether these settled materials might release soluble toxic species (e.g. in the case of nano-silver) or if isolation of settled materials

Section	Key research topics
	precludes potential physical effects of some MNs (e.g. as reported by [15]). • Research could also focus more broadly on quantifying the uncertainties that arise when exposure varies beyond specific thresholds (including ± 20%).
Dosimetry and Interpretation	• It is important to continue to develop, validate and standardize analytical methods to assess particle number concentrations and size distributions in aqueous samples. As described above, spICP-MS is one method that shows promise in this regard but the practicality of using this approach in routine hazard testing is unclear. • Research is needed to develop analytical techniques and reference material standards to directly measure total available surface area and particle number concentration at toxicologically relevant concentrations (sometimes low µg L^{-1}). Without readily available direct measurement methods, it will be difficult to relate dose response to surface area or particle number metrics for heterodispersed suspensions of MNs that are unstable in biological media over time. • The impact of consistent versus inconsistent concentration, agglomeration and dissolution on ecotoxicity test results needs to be assessed; ideally, each of these factors for a relevant test material could be isolated and tested. It would be preferable to have a consistent particle and environmentally relevant dosage but the feasibility of accomplishing this is unclear. • Lastly, an assessment of the robustness of alternative dose metrics is necessary. This could be accomplished by assessing to what extent the uncertainty in the measurements of alternative dose metrics would influence the uncertainty of the final results. It may be that the accuracy of toxicity endpoint values for the alternative dose metrics is highly limited by the great variability of the measurements used to determine dose metrics in aqueous media and the extent to which the exposure changes during the bioassay.
Sediment testing	• A combined research and regulatory need is to determine when it is or is not appropriate to conduct a sediment toxicity test for a MN, perhaps based on the material behavior in the water column and interaction with sediment ligands. A decision tree based on particle behavior, hazard and a series of functional tests may be a logical approach. • Additional needs clearly involve development of characterization methods for particles in the sediment matrix. This is particularly important for measurements of carbon-based MNs which cannot be readily measured using element specific methods. For metal and metal oxide MNs, the development of methods to differentiate between MNs, dissolved metal ions, and MN agglomerates is needed. • In addition, the reliability and potential for creating homogeneous MN distributions in sediments or on their surfaces using the different MN spiking methods should be evaluated. • The reliability of porewater measurements should also be evaluated. If these methods utilize centrifugation to remove sediment particles, these methods may also potentially remove MNs. If that occurs, then new methods for making porewater measurements are needed or these measurements would only measure freely dissolved ions for MNs that readily dissolve.

References

1. OECD, *Guidance Document on Transformation/Dissolution of Metals and Metal Compounds in Aqueous Media*. OECD Publishing: 2002.
2. APHA (American Public Health Association). *Standard method for the examination of water and waste water*. 19 ed.; APHA: Washington, DC, 1995.
3. EPA (Environmental Protection Agency). *Monitoring Trace Metals at Ambient Water Quality Criteria Levels*; EPA: Washington, DC, 1995.
4. EPA (Environmental Protection Agency). *Trace Metal Clean Sampling of Natural Waters*; EPA: Richmond, CA, 2004.
5. Connell, D. W., Basic Concepts of Environmental Chemistry. CRC Press/Lewis: Boca Raton, FL, 1997.
6. Pakarinen, K.; Petersen, E. J.; Alvila, L.; Waissi-Leinonen, G. C.; Akkanen, J.; Leppanen, M. T.; Kukkonen, J. V. A screening study on the fate of fullerenes (nC60) and their toxic implications in natural freshwaters. *Environ. Toxicol. Chem.* **2013,** *32* (6), 1224-32.
7. Akkanen, J.; Penttinen, S.; Haitzer, M.; Kukkonen, J. V. Bioavailability of atrazine, pyrene and benzo[a]pyrene in European river waters. *Chemosphere* **2001,** *45* (4-5), 453-62.
8. OECD, (Organisation for Economic Co-operation and Development). *Guidance Document on Aquatic Toxicity Testing of Difficult Substances and Mixtures*; OECD: Paris, France, 2000.
9. OECD, *Test No. 201: Freshwater Alga and Cyanobacteria, Growth Inhibition Test*. OECD: Paris, France, 2011.
10. OECD (Organisation for Economic Co-operation and Development). *Test No. 211: Daphnia magna Reproduction Test*. OECD: Paris, France, 2012.
11. Simpson, S. L.; Roland, M. G.; Stauber, J. L.; Batley, G. E. Effect of declining toxicant concentrations on algal bioassay endpoints. *Environ. Toxicol. Chem.* **2003,** *22* (9), 2073-2079.
12. Palmqvist, A.; Baker, L.; Forbes, V.; Gergs, A.; von der Kammer, F.; Luoma, S.; Lützhøft, H.; Salinas, E.; Sorensen, M.; Steevens, J., Nanomaterial environmental risk assessment. *Integrated Environ. Assess. Manag.* **2015,** *11* (2), 333-335.
13. Petersen, E. J.; Henry, T. B.; Zhao, J.; MacCuspie, R. I.; Kirschling, T. L.; Dobrovolskaia, M. A.; Hackley, V.; Xing, B.; White, J. C. Identification and avoidance of potential artifacts and misinterpretations in nanomaterial ecotoxicity measurements. *Environ. Sci. Technol.* **2014,** *48* (8), 4226-4246.
14. Hartmann, N. B.; Engelbrekt, C.; Zhang, J.; Ulstrup, J.; Kusk, K. O.; Baun, A. The challenges of testing metal and metal oxide nanoparticles in algal bioassays: titanium dioxide and gold nanoparticles as case studies. *Nanotoxicology* **2013,** *7* (6), 1082-94.
15. Dabrunz, A.; Duester, L.; Prasse, C.; Seitz, F.; Rosenfeldt, R.; Schilde, C.; Schaumann, G. E.; Schulz, R. Biological surface coating and molting inhibition as mechanisms of TiO_2 nanoparticle toxicity in *Daphnia magna*. *PLoS ONE* **2011,** *6* (5), e20112.