Organisation for Economic Co-operation and Development

ENV/JM/MONO(2019)13

Unclassified

English - Or. English

27 May 2019

ENVIRONMENT DIRECTORATE
JOINT MEETING OF THE CHEMICALS COMMITTEE AND THE WORKING PARTY ON CHEMICALS, PESTICIDES AND BIOTECHNOLOGY

GUIDING PRINCIPLES FOR MEASUREMENTS AND REPORTING FOR NANOMATERIALS: PHYSICAL CHEMICAL PARAMETERS

Series on the Safety of Manufactured Nanomaterials
No. 91

JT03448011

OECD Environment, Health and Safety Publications

Series on the Safety of Manufactured Nanomaterials

No. 91

GUIDING PRINCIPLES FOR MEASUREMENTS AND REPORTING FOR NANOMATERIALS: PHYSICAL CHEMICAL PARAMETERS

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

A cooperative agreement among FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris, 2019

Also published in the Series on the Safety of Manufactured Nanomaterials:

No. 1, *Report of the OECD Workshop on the Safety of Manufactured Nanomaterials: Building Co-operation, Co-ordination and Communication (2006)*

No. 2, *Current Developments/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 1st Meeting of the Working Party on Manufactured Nanomaterials (2006)*

No. 3, *Current Developments/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 2nd Meeting of the Working Party on Manufactured Nanomaterials (2007)*

No. 4, *Manufactured Nanomaterials: Programme of Work 2006-2008 (2008)*

No. 5, *Current Developments/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 3rd Meeting of the Working Party on Manufactured Nanomaterials (2008)*

No. 6, *List of Manufactured Nanomaterials and List of Endpoints for Phase One of the OECD Testing Programme (2008)*

No. 7, *Current Developments/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 4th Meeting of the Working Party on Manufactured Nanomaterials (2008)*

No. 8, *Preliminary Analysis of Exposure Measurement and Exposure Mitigation in Occupational Settings: Manufactured Nanomaterials (2009)*

No. 9, *EHS Research Strategies on Manufactured Nanomaterials: Compilation of Outputs (2009)*

No. 10, *Identification, Compilation and Analysis of Guidance Information for Exposure Measurement and Exposure Mitigation: Manufactured Nanomaterials (2009)*

No. 11, *Emission Assessment for the Identification of Sources and Release of Airborne Manufactured Nanomaterials in the Workplace: Compilation of Existing Guidance (2009)*

No. 12, *Comparison of Guidance on Selection of Skin Protective Equipment and Respirators for Use in the Workplace: Manufactured Nanomaterials (2009)*

No. 13, *Report of an OECD Workshop on Exposure Assessment and Exposure Mitigation: Manufactured Nanomaterials (2009)*

No. 14, *Guidance Manual for the Testing of Manufactured Nanomaterials: OECD Sponsorship Programme (2009)*

No. 15, *Preliminary Review of OECD Test Guidelines for their Applicability to Manufactured Nanomaterials (2009)*

No. 16, *Manufactured Nanomaterials: Work Programme 2009-2012 (2009)*

No. 17, *Current Development/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 5th Meeting of the Working Party on Manufactured Nanomaterials (2009)*

No. 18, *Manufactured Nanomaterials: Roadmap for Activities during 2009 and 2010 (2009)*

No. 19, *Analysis of Information Gathering Initiative on Manufactured Nanomaterials (2009)*

No. 20, *Current Development/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 6th Meeting of the Working Party on Manufactured Nanomaterials (2010)*

No. 21, *Report of the Workshop on Risk Assessment of Manufactured Nanomaterials in a Regulatory Context (2010)*

No. 22, *OECD Programme on the Safety of Manufactured Nanomaterials 2009-2012: Operational Plans of the Projects (2010)*

No. 23, *Report of the Questionnaire on Regulatory Regimes for Manufactured Nanomaterials (2010)*

No. 24, *Preliminary Guidance Notes on Sample Preparation and Dosimetry for the Safety Testing of Manufactured Nanomaterials (2010)*

No. 25, *Guidance Manual for the Testing of Manufactured Nanomaterials: OECD Sponsorship Programme: First Revision (2010)*

No. 26, *Current Development/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 7th Meeting of the Working Party on Manufactured Nanomaterials (2010)*

No. 27, *List of Manufactured Nanomaterials and List of Endpoints for Phase One of the Sponsorship Programme for the Testing Manufactured Nanomaterials: Revised (2010)*

No. 28, *Compilation and Comparison of Guidelines Related to Exposure to Nanomaterials in Laboratories (2010)*

No. 29, *Current Development/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 8th Meeting of the Working Party on Manufactured Nanomaterials (2011)*

No. 30, *Regulated Nanomaterials: 2006-2009(2011)*

No. 31, *Information Gathering Schemes on Nanomaterials: Lessons Learned and Reported Information (2011)*

No. 32, *National Activities on Life Cycle Assessment of Nanomaterials (2011)*

No. 33, *Important Issues on Risk Assessment of Manufactured Nanomaterials (2012)*

No. 34, *Current Development/ Activities on the Safety of Manufactured Nanomaterials: Tour de table at the 9th Meeting of the Working Party on Manufactured Nanomaterials (2012)*

No. 35, *Inhalation Toxicity Testing: Expert Meeting on Potential Revisions to OECD Test Guidelines and Guidance Document (2012)*

No. 36, *Guidance on Sample Preparation and Dosimetry for the Safety Testing of Manufactured Nanomaterials (2012)*

No.37, *Current Developments in Delegations on the Safety of Manufactured Nanomaterials - Tour de Table at the 10th Meeting of the WPMN (2012)*

No.38, *Co-Operation on Risk Assessment: Prioritisation of Important Issues on Risk Assessment of Manufactured Nanomaterials - Final Report (2013)*

No. 39, *Environmentally Sustainable Use of Manufactured Nanomaterials - Workshop Report (2013)*

No. 40, *Ecotoxicology and Environmental Fate of Manufactured Nanomaterials: Test Guidelines (2014)*

No.41, *Report of the OECD Expert meeting on the Physical Chemical Properties of Manufactured Nanomaterials and Test Guidelines (2014)*

No.42, *Report of the questionnaire on regulatory regimes for manufactured nanomaterials 2010-2011 (2014)*

No.43, *Genotoxicity of Manufactured Nanomaterials: Report of the OECD expert meeting (2014)*

Nos. 44-54, These items are the dossiers derived from the Testing Programme on Manufactured Nanomaterials, which are located at:

http://www.oecd.org/chemicalsafety/nanosafety/testing-programme-manufactured-nanomaterials.htm

No.55, *Harmonized Tiered Approach to Measure and Assess the Potential Exposure to Airbone Emissions of Engineered Nano-objects and their Agglomerates and Aggregates at Workplaces (2015)*

No.56, *Analysis of the Survey on Available Methods and Models for Assessing Exposure to Manufactured Nanomaterials (2015)*

No.57, *Guidance Manual towards the integration of risk assessment into life cycle assessment of nano-enabled applications (2015)*

No.58, *Preliminary guidance notes on Nanomaterials: Interspecies variability factors in human health risk assessment (2015)*

No.59, *Developments on the safety of manufactured nanomaterials: 2013 (2015)*

No.60, *Current developments in delegations on the safety of manufactured nanomaterials - tour de table (2015)*

No.61, *Developments in delegations on the safety of manufactured nanomaterials - tour de table (2015)*

No.62, *Considerations for using dissolution as a function of surface chemistry to Evaluate environmental behaviour of nanomaterials in risk assessments (2015)*

No.63, *Physical-chemical parameters: measurements and methods relevant for the regulation of nanomaterials (2016)*

No.64, *Approaches on nano grouping/ equivalence/ read-across concepts based on physical-chemical properties (GERA-PC) for regulatory regimes (2016)*

No.65, *Physical-chemical properties of nanomaterials: Evaluation of methods applied in the OECD-WPMN testing programme (2016)*

No.66, *Categorisation of manufactured nanomaterials (2016)*

No.67, *Developments in delegations on the safety of manufactured nanomaterials - tour de table (2016)*

No.68, *Multi-walled carbon nanotubes (MWCNT): summary of the dossier (2016)*

No.69, *Fullerenes (C60): summary of the dossier (2016)*

No.70, *Single walled carbon nanotubes (SWCNTs): summary of the dossier (2016)*

No.71, *Silicon dioxide: summary of the dossier (2016)*

No.72, *Toxicokinetics of manufactured nanomaterials: report from the OECD expert meeting (2016)*

No.73, *Titanium dioxide: summary of the dossier (2016)*

No.74, *Exposure Assessment of Nano-Silver (AgNP): Case Study (2016)*

No.75, *Future Challenges Related to the Safety of Manufactured Nanomaterials: Report from the Special Session (2016)*

No.76, *Grouping and Read-Across for the Hazard Assessment of Manufactured Nanomaterials: Report from the Expert Meeting (2016)*

No.77, *Gold Nanoparticle Occupational Exposure Assessment in a Pilot Scale Facility: Nanomaterials Exposure Case Study (2016)*

No.78, *Developments on the Safety of Manufactured Nanomaterials: Tour de Table from OECD Delegations (Nov.2015 - Oct. 2016)*

No.79, *Strategy for Using Metal Impurities as Carbon Nanotube Tracers (2016)*

No.80, *Alternative testing strategies in risk assessment of manufactured nanomaterials: current state of knowledge and research needs to advance their use (2017)*

No.81, *Developments in Delegations on the Safety of Manufactured Nanomaterials – Tour de Table (2017)*

No. 82, *Strategies, Techniques and Sampling Protocols for Determining the Concentrations of Manufactured Nanomaterials in Air at the Workplace (2017)*

No. 83, *Silver Nanoparticles: Summary of the Dossier (2017)*

No. 84, *Consumer and environmental exposure to manufactured nanomaterials - Information used to characterize exposures: Analysis of a Survey (2017)*

No. 85, *Evaluation of in Vitro Methods for Human Hazard Assessment Applied in the OECD Testing Programme for the Safety of Manufactured Nanomaterials (2018)*

No. 86, *Assessment of Biodurability of Nanomaterials and their Surface Ligands (2018)*

No. 87, *Developments in Delegations on the Safety of Manufactured Nanomaterials - Tour de Table (2018)*

No. 88, *Investigating the Different Types of Risk Assessments of Manufactured Nanomaterials*

No. 89, *Developments in Delegations on the Safety of Manufactured Nanomaterials - Tour de Table (2019)*

No. 90, *Physical-Chemical Decision Framework to Inform Decisions for Risk Assessment of Manufactured Nanomaterials*

ABOUT THE OECD

The Organisation for Economic Co-operation and Development (OECD) is an intergovernmental organisation in which representatives of 36 industrialised countries in North and South America, Europe and the Asia and Pacific region, as well as the European Commission, meet to co-ordinate and harmonise policies, discuss issues of mutual concern, and work together to respond to international problems. Most of the OECD's work is carried out by more than 200 specialised committees and working groups composed of member country delegates. Observers from several countries with special status at the OECD, and from interested international organisations, attend many of the OECD's workshops and other meetings. Committees and working groups are served by the OECD Secretariat, located in Paris, France, which is organised into directorates and divisions.

The Environment, Health and Safety Division publishes free-of-charge documents in 12 different series: **Testing and Assessment; Good Laboratory Practice and Compliance Monitoring; Pesticides; Biocides; Risk Management; Harmonisation of Regulatory Oversight in Biotechnology; Safety of Novel Foods and Feeds; Chemical Accidents; Pollutant Release and Transfer Registers; Emission Scenario Documents; Safety of Manufactured Nanomaterials,** and **Adverse Outcome Pathways**. More information about the Environment, Health and Safety Programme and EHS publications is available on the OECD's World Wide Web site (www.oecd.org/chemicalsafety/).

This publication was developed in the IOMC context. The contents do not necessarily reflect the views or stated policies of individual IOMC Participating Organisations.

The Inter-Organisation Programme for the Sound Management of Chemicals (IOMC) was established in 1995 following recommendations made by the 1992 UN Conference on Environment and Development to strengthen co-operation and increase international co-ordination in the field of chemical safety. The Participating Organisations are FAO, ILO, UNDP, UNEP, UNIDO, UNITAR, WHO, World Bank and OECD. The purpose of the IOMC is to promote co-ordination of the policies and activities pursued by the Participating Organisations, jointly or separately, to achieve the sound management of chemicals in relation to human health and the environment.

This publication is available electronically, at no charge.

For this and many other Environment, Health and Safety publications, consult the OECD's World Wide Web site (www.oecd.org/chemicalsafety/)

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2 rue André-Pascal
75775 Paris Cedex 16
France

Fax: (33-1) 44 30 61 80

E-mail: ehscont@oecd.org

FOREWORD

The OECD Working Party on Manufactured Nanomaterials (WPMN) is a subsidiary body of the OECD Chemicals Committee. This programme concentrates on human health and environmental safety implications of manufactured nanomaterials (limited mainly to the chemicals sector), and aims to ensure that the approach to hazard, exposure and risk assessment is of a high, science-based, and internationally harmonised standard. It promotes international co-operation on the human health and environmental safety of manufactured nanomaterials, and involves the safety testing and risk assessment of manufactured nanomaterials.

Physico-chemical properties are key starting points for risk assessments of chemicals. They provide a description of the chemical, and prove useful in assessment of environmental behaviour, uptake routes into organisms, toxicokinetics and ultimate effects in organisms. For nanomaterials, relevant physico-chemical properties can differ from those commonly considered for non-nanomaterials (e.g. surface area versus boiling point). With this in mind, The OECD WPMN has striven to develop tools that can assist identifying appropriate physico-chemical parameters to better understand the link between those parameters and potential human health and environmental effects of nanomaterials, and to facilitate prediction of such effects. This document supports the Physical-Chemical Decision Framework To Inform Decisions For Risk Assessment Of Manufactured Nanomaterials [ENV/JM/MONO(2019)12]. It is intended to aid in improving the conduction of the studies, in addition to promoting consistent data reporting (including reporting details on sample preparation and measurement protocols) to maximise utility and comparability of the data. These two project are complementary and are collectively intended to facilitate the identification of the most useful parameters and best available methods while maintaining rigour in data quality and reporting. It is important to note that there are not intended for risk assessment per se but rather presents an approach to gather fit-for-purpose physico-chemical information to more fully understand the behaviour of nanomaterials in biotic and abiotic systems.

This framework, and its guiding principles, are recognised as a living document, and they will be subject to amendment and refinement as researchers gain greater understanding in using it.

This document is published under the responsibility of the Joint Meeting of the Chemicals Committee and the Working Party on Chemicals, Pesticides and Biotechnology of the OECD.

Table of contents

Tables

Figures

Introduction

1. Fit-for-purpose testing and data reporting are key components to regulatory decision making. For many methods, OECD test guidelines exist, providing transparent, tested, and broadly accepted measurement procedures for the generation of specific and comparable data. This data is accepted globally through the Mutual Acceptance of Data (MAD) principle[1]. The MAD principle greatly reduces ambiguity in addressing requirements, the research burden and the resources required for evaluating chemical substances. However, OECD and other accepted test guidelines are not available for each and every potential physico-chemical endpoint (e.g. surface reactivity, hetero-coagulation). This necessitates the use of information from non-standardized methods – often not conducted under the conditions of Good Laboratory Practices (GLP) – for information gathering or for the assessment of substances.

2. Currently, data generated by methods not covered by MAD does not have an associated consistent manner to treat and evaluate it. Regional differences in information requirements and intended use of the information further complicates the question of data and method suitability, even for the same general physico-chemical parameter. One example is that a suitable data set with associated method(s) for determining the particle size distribution of a substance may vary from region to region. Thus to share data, communication becomes extremely important, especially when sharing on a global level. In addition, any substance and local environmental scenario specific testing adaptations (due to the extrinsic nature of many potentially relevant physico-chemical parameters) require improved communication practices in order to communicate important alterations in the test systems with supporting rationale. This not only aids in appropriately interpreting and comparing (or not comparing) outcomes, but also aids in the development of an improved testing strategy. To further this along, a common process for method selection and data quality assessment would be beneficial when clear guidelines do not exist.

3. Physico-chemical properties are considered essential parameters and a starting point for any risk assessment, including the grouping of traditional chemical substances and nanomaterials. A lack of both standard measurement methods and sample preparation protocols has limited progress in developing relationships between physico-chemical parameters and aspects related to the health and safety of nanomaterials.

4. Physico-chemical characterisation endpoints can be measured through multiple methods, each with their own nuances and insights into the material's behaviour and properties. Varied sample preparation requirements and approaches often accompany these measurement methods, and, frequently, sample preparation and measurement protocols require adjustment in order to provide the most relevant information for different purposes (e.g. for material identification versus informing exposure potential). Nevertheless, the general process for determining method suitability, limitations, and necessary reporting requirements for comparability with other techniques/equipment is similar.

5. For hazard assessment, a set of principles were developed and agreed upon at an OECD Workshop in Solna, Sweden in 1996, for the validation of new or updated test methods

[1] www.oecd.org/env/ehs/mutualacceptanceofdatamad.htm

for hazard assessment. These "Solna Principles" though are also applicable to the physico-chemical testing of nanomaterials and provide a background philosophy that is core to the objectives of this document. The Principles are:

1) *A rationale for the test method should be available. This should include a clear statement of scientific need and regulatory purpose.*
2) *The relationship of the endpoint(s) determined by the test method to the in vivo biological effect and to the toxicity of interest should be addressed. The limitations of a method should be described, e.g., metabolic capability.*
3) *A formal detailed protocol must be provided and should be readily available in the public domain. It should be sufficiently detailed to enable the user to adhere to it, and it should include data analysis and decision criteria. Test methods and results should be available preferably in an independent peer-reviewed publication. In addition, the result of the test should have been subjected to independent scientific review.*
4) *Intra-test variability, repeatability and reproducibility of the test method within and amongst laboratories should have been demonstrated. Data should be provided describing the level of inter- and intra-laboratory variability and how these vary with time.*
5) *The test method's performance must have been demonstrated using a series of reference chemicals preferably coded to exclude bias.*
6) *The performance of test methods should have been evaluated in relation to existing relevant toxicity data as well as information from the relevant target species.*
7) *All data supporting the assessment of the validity of the test methods including the full data set collected in the validation study must be available for review.*
8) *Normally, these data should have been obtained in accordance with the OECD Principles of Good Laboratory Practice (GLP).*

6. The Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment (ENV/JM/MONO(2005)14) also provides detailed information on the conduction and validation of test methods for hazard assessment. This document provided detailed information on study design, different approaches to validation, important supporting documentation for new test submissions and general criteria for regulatory acceptance. As with the Solna Principles, the fundamental approaches and considerations are also relevant for the physico-chemical testing of nanomaterials. To this regard, in addition to the Solna Principles, the following criteria are also important to the regulatory acceptance of a test method or testing approach:

1) *Application of the method provides data that adequately predicts the end-point of interest in that it demonstrates either a linkage between (i) the new test and an existing test method or (ii) the new test and effects in the target species.*
2) *The method generates data for risk assessment purposes that are at least as useful as, and preferably better than, those obtained using existing methods. This will give a comparable or better level of protection for human health or the environment.*

3) There are adequate testing data for chemicals and products representative of the type of chemicals administered by the regulatory programme or agency (e.g. pesticides, cosmetics).

4) The test should be robust and transferable and allow for standardisation. If highly specialised equipment, materials or expertise are required, efforts should be sought to facilitate transferability. This is an important criterion to be considered at an early stage of a validation study. [Note added by the Secretariat: According to current OECD policy, the test should not require equipment or material from a unique source. This would prevent the acceptance of patented methods. The Solna Workshop did not discuss the issue of patented tests but referred the issue to higher policy levels at OECD].

5) The test is cost effective and likely to be used.

6) Justification (scientific, ethical, economic) should be provided for the new method with respect to any existing methods available. In this respect due consideration should be given to animal welfare consideration including the 3Rs.

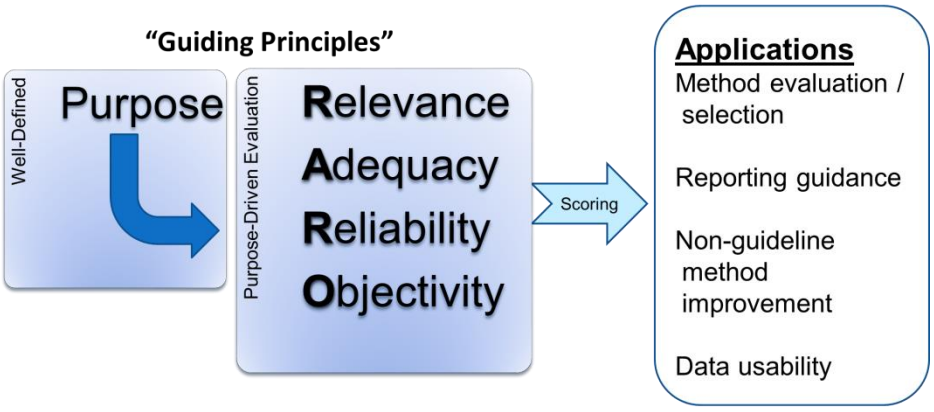
7. This document builds upon the concepts of the Solna Principles and the Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment (ENV/JM/MONO(2005)14) and provides adaptations for the physico-chemical characterization of nanomaterials. However, the current document is not intended to guide the development of OECD Test Methods, but rather to provide a means of identifying fit-for-purpose and suitable quality methods from a number of sources – some of which may not have been intended for regulatory purpose (e.g. academic research). The present document also recognizes that the level of acceptable precision and rigour varies for different purposes of physico-chemical characterization, not all of which may be related to hazard or exposure, but simply for identification. Parameters important for identifying a substance but not linked to hazard or exposure potential may not require as detailed scrutiny as those that have shown a clear link.

8. This document provides a transparent approach that (1) aids communication of key purposes for the data generation, (2) facilitates the identification of suitable methods, (3) pinpoints method limitations and (4) highlights good reporting practices that address purposes related to the assessment of nanomaterials. Taken as a whole this transparent approach is called "the guiding principles". This document introduces their usage for different types of physico-chemical analysis. The approach provided within this document has been developed to accompany the Physico-Chemical Decision Framework to Inform Decisions for Risk Assessment and integrates the concepts of the Solna Principles for physico-chemical characterization.

Guiding Principles: Overview, Parameters and Descriptions

9. Overall, the guiding principles are intended to afford a systematic and objective system for identifying suitable analytical methods, data sources and technical approaches to address knowledge gaps related to the safety assessment of nanomaterials. The applicability of many of these principles is broader than nanomaterials. The Guiding Principles have been developed incorporating the learnings from OECD GD 34 (ENV/JM/MONO(2005)14). It integrates and builds on several established quality assessment and strategic experimentation methods (e.g. Six Sigma Methodology) that have been applied for industrial, analytical and regulatory purposes. For the guiding principles, the clear communication of the intended purpose of the measurement and explicit requirements from the requestor of the data (e.g. regulatory organisation) are important to determine the fit-for-purpose status of evaluated methodologies and data sets. The approach is also intended to provide a transparent and structured evaluation tool (based on common quality principles) to facilitate ranking and selection of methods, protocols, or data sources. This assists in clarifying the deficiencies and benefits of different method options with respect to the identified purpose. The individual guiding principles highlighted in Figure 1are discussed below, followed by sections describing their practical implementation.

Figure 1. Schematic identifying the Guiding Principles and potential applications



Note: *OECD GD 34 provides additional detail on requirements for "Relevance" and "Reliability."*

10. Collectively, the Guiding Principles address key areas of concern when attempting to use, apply, or recommend physico-chemical methods and data for the evaluation of nanomaterials and, more generally, chemical substances. The guiding principles are summarised in Table 1.

Table 1. List and Description of the Guiding Principles for Measurement and Reporting of Nanomaterials

Guiding Principle	Description
Purpose	A clear description of the need and intended use of the data with defined domain of applicability. For example, measurement of particle size distribution for the assessment of substance identity or for the assessment of environmental fate. Purposes with associated physico-chemical endpoints/parameters are defined in the Decision Framework
Relevance	A measure of the degree of alignment of the overall methods and/or data with regards to the intended purpose. Methodology capable of identifying key mechanism should be prioritised, if possible.
Reliability	A measure of the repeatability, reproducibility, trueness (accuracy) and suitability of the reported data and applied methodologies with respect to necessary degree of resolution as specified by the intended purpose.
Adequacy	A measure of the completeness of the methodological description and reported data with respect to the intended purpose.
Objectivity	A measure of the extent of bias due to sampling, estimations and systematic effects based on the overall methods employed and study design with respect to the intended purpose.

11. In the following sections, each guiding principle is described in detail with respect to parameters relevant for method selection, data acceptance, and data reporting. Subsequently, their integration into a comprehensive assessment tool is discussed, followed by a discussion of the integration of the guiding principles with the Physico-Chemical Decision Framework to Inform Decisions for Risk Assessment and then exemplified in case studies.

Purpose

12. Communication of the purpose for the data generation should, for the sake of clarity, be supported by meta-data, i.e. context setting parameters. These include the identified information need, intended use of the information, relevant substances (i.e. is the data to fulfil this query specific to a subset of nanomaterial types or particle types), and the actual scenario that the information is intended to aid in predicting. For instance, a request for the dissolution rate of a coated particle substance in lung lining fluid should provide additional descriptive information. Knowing how the test is intended to be used and which substances are to be represented by the test is critical. If the data was intended to draw a detailed comparison to another particle type with a known *in vivo* dissolution profile, then the test would be designed differently (and assessed for quality differently) than if the data was only intended for rough screening applying a standard media to determine if more rigorous assessment is needed.

13. It is noted that the companion project, the Physico-Chemical Decision Framework, can be used to identify purposes and the desired physico-chemical information. This document is intended to aid in identifying the suitability of the methods selected. For instance, if a particle size distribution is intended for the identification of a nanomaterial, a primary particle size distribution may be suitable with appropriate validation. However, if the particle size distribution is intended to determine the potential for pulmonary exposure an aerosolized particle size distribution (with appropriate reference measurement and validation) or dustiness measurements may be

desired as an indication of inhalation exposure potential. Although qualitative, this information is helpful for identifying best available options (e.g., methods, data sets), when an ideal solution may not be available and choices between suboptimal information gathering is needed. For instance, determining the surface reactivity of a nanomaterial after release into the environment would be complicated by a wide range of factors including surrounding media and the preceding environment/matrix from which it was released. The Guiding Principles can then be used to select the most appropriate method for the purpose of measuring the surface reactivity of the nanomaterial as released in a specific environment.

14. Additionally, quantitative parameters are also important to identify if essential requirements are met, in addition to ensuring that the applied method/data resolution and accuracy are suitable for the given purpose. Depending on purpose, pre-established requirements in the form of critical-to-quality (CTQ) elements may need to be addressed. CTQs vary in form but are generally used to highlight pre-identified critical factors that are necessary to be fit-for-purpose. CTQs can range from mandating sample compatibility for developing methodological cross-correlations to facilitate data inter-exchangeability between historical benchmark methods and proposed methods, to simply meeting cost and qualitative expertise requirements. For methods that might be considered for OECD a CTQ might be confirmation that the method is not restricted to proprietary instrumentation or reagents, is reasonably available and cost effective. An understanding of how precise the data is required to be in order to resolve meaningful differences is another key factor. This is separate from accuracy requirements that identify how close to true values the data must be, with the resolution limits of the methodology in mind.

15. The identification and clear communication of the purpose is essential for all subsequent guiding principles. Hence, clearly defining the purpose is critical. For a purpose to be meaningful there needs to be enough specificity to enable the identification of methods that are clearly fit-for-purpose and those that are clearly not fit-for-purpose. In terms of physico-chemical parameters, the purpose for a given measurand can vary widely (e.g. dissolution measurements for determining solubility in fresh water systems versus in lung lining fluid). Sufficient detail and context including both informative and quantitative parameters is required to facilitate practical fit-for-purpose measurements and reporting. Table 2 provides the relevant parameters for defining Purpose.

Table 2. Considerations regarding Purpose in order to evaluate the requirements to information generation

Parameters	Related Questions
Information Requirement	Why is the information required?
Intended Use	How will the information be used? How will the information not be used?
Substance Identification	For which substance(s) will the information apply? Which parameters identify similar and dissimilar substances for the evaluated material class?
Relevant Realistic Scenarios	What realistic scenarios would the data aid in predicting? Is the information intended to be broadly or narrowly applied? What are the relevant conditions that need to be satisfied?
Critical-to-quality (CTQ) Components	Are there specific key factors, related to the purpose, implicated as critical to meeting the overall purpose? (e.g. method is cost effective and widely available)
Resolution Requirement	What is the required resolution of the data? (e.g. the practical quantification difference to be made against another substance – how small of a difference must be quantified). A precision specification (e.g. within +/- 10%) should be given, if relevant.
Accuracy and Range Requirement	How accurate (true) is the data required to be and over what quantification range? Identify relevant reference standards, gold standard methods (if any).

16. Throughout this document, considerations for Purpose are reflected in order to determine relevance, adequacy, reliability and objectivity. Hence, it is critical that the Purpose is linked to a clear description of the need and intended use of the data with defined domain of applicability, also to ensure that the subsequent evaluation will be useful. The parameters and questions listed in Table 2 are intended to ensure that both the broader and specific context for the data is appropriately addressed. By identifying core elements within the other guiding principles that need to be met to satisfy the purpose, a fit-for-purpose status of the methods can be determined. Further specification and refinement of the Purpose through communication from/with the data accessor would be needed in case that there is insufficient information available to clearly determine whether the purpose has been reasonably satisfied. The *Integrated Assessment and Scoring Protocol* section of this document suggests and describes a process for assessing this.

Relevance

17. Relevance is a measure of the degree of alignment of the overall methods and/or data in regards to the intended purpose. Some assessments refer to this measure as directness. It includes specification of the relevance of the scientific principle, the test method, the applied media, sample preparation and sample history with respect to the intended purpose. Parameters related to relevance, many specific to the physico-chemical characterisation of nanomaterials, are provided in Table 3.

Table 3. Considerations regarding Evaluating Relevance of available information

Parameters	Description
Scientific Principle	The degree to which the applied scientific principle used in data acquisition fits the intended purpose. Is the measurement capable of measuring the desired parameter conclusively without broad assumptions?
Substance	The degree of similarity of the substance analysed and the intended substance(s) identified in the purpose. Does the measured substance represent the intended substance?
Media & Environmental Conditions	The degree of appropriateness of the test media (i.e. liquid, solid, and gaseous matrix used in the test) and environmental conditions (e.g. temperature, humidity) for the intended purpose.
Sample Preparation	The suitability of the sample preparation methodology for the intended purpose.
Sample Concentration	The appropriateness of the sample concentration with respect to the realistic phenomena identified in the purpose.
Time	The appropriateness of the timescale over which the experiment was conducted with respect to the realistic phenomena identified in the purpose.
History	The appropriateness of the sample storage, handling, and conditioning prior to testing (i.e., sample history) with respect to the realistic experience for the phenomena intended to be mimicked as identified in the purpose (e.g. for materials that may transform, is the appropriate form of the substance being tested?)
Critical-to-quality (CTQ) Compliance	The degree to which the data or methodology is within an acceptable range facilitating compliance with all CTQs.

Adequacy

18. Adequacy is a measure of the completeness of the methodological description and reported data with respect to the intended purpose. The evaluation of the adequacy of test results and documentation is particularly important when the techniques and analysis methods are uncommon. Parameters related to adequacy are presented in Table 4.

Table 4. Considerations regarding Evaluating Adequacy of reporting

Parameters	Description
Analytical Process Description	The degree to which the analytical process description allows for one to independently reproduce the experiment with certainty. This includes sufficient descriptions of environmental conditions, applied media, sample preparation methodology, and measurement timescales.
Substance Description	The thoroughness of the description of the substance origin (e.g. lot #, date of manufacture, synthesis route).
Storage and Handling Description	The thoroughness of the description of the storage and sequence of handling of the substance after initial receipt and prior to analysis.
Description of sources of experimental error and uncertainty	The degree to which source of experimental errors and uncertainty have been contemplated and conveyed.

Reliability

19. Reliability is measure of the reproducibility of the reported data from the applied methodologies with respect to necessary degree of resolution as specified by the intended purpose. Reliability includes validation factors specific to the analytical technique as well as laboratory procedure. These parameters specified are typically addressed during method development and the determination of standard laboratory protocols (SOPs). When SOPs have been developed these parameters should be noted within the associated documentation. The parameters associated with reliability are presented in Table 5.

Table 5. Considerations regarding Evaluating Reliability of the method

Parameters	Description
Technique Related:	
Accuracy	The degree at which the technique meets the specified range for accuracy. The accuracy of an analytical procedure expresses the closeness of agreement between the value which is accepted either as a conventional true value or an accepted reference value and the value found.
Specificity	The ability of the method to measure the desired parameter unequivocally considering likely confounding factors.
Limit of Quantification (LOQ)	The degree in which the lowest amount of a parameter that can be quantitatively determined (with suitable precision and accuracy) fits within the resolution requirement.
Linearity	The degree to which the method produces a result directly proportional to the quantity of a parameter over the range of the technique.
Range	The degree to which the range of the technique matches the desired range. The range of an analytical procedure is the interval between the upper and lower amounts of a parameter for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.
Robustness	The degree to which the technique is unaffected by small, but deliberate variations in method parameters. (An indicator of reliability under normal usage)
Laboratory Related:	
Operator Training	The extent to which the laboratory applies formal validation protocols for the qualification of technique operators.
Instrument Validation	The extent to which the laboratory implements regular instrument calibrations and controls to identify deviations in performance.
Media Characterization	The extent to which media applied in the testing has been characterised.
Reference Measurements	The extent to which reference material measurements have been applied to evaluate accuracy and reproducibility and to aid in normalising results.
Intermediate precision	The degree of within laboratory variations in reported outcomes amongst available equipment and operators.
Reproducibility	The degree to which laboratory practices for the parameter have been confirmed via comparison with results from external laboratories. (e.g. participation in inter-laboratory comparison / round robin testing).

Overall Laboratory Experience Level	When appropriate, the degree of familiarity of laboratory staff to the technique and its fundamental principles.

Objectivity

20. Objectivity is a measure of the extent of error due to sampling, estimations and systematic effects based on the overall methods employed and study design with respect to the intended purpose. The key dimensions for objectivity are Bias and Transparency as indicated in Table 6 adapted from the Cochrane Collaboration's Risk of Bias Tool (Higgins et al. 2011) and Bioletta et al. (2014). Identifying potential sources of bias is an important step in qualifying data and methodologies. Relevant parameters of bias are given in Table 7. Notably, to judge bias, a suitable amount of transparency is required. An increasing number of peer-reviewed journals are currently adopting transparency measures to improve the utility and quality of research. Pertinent parameters of transparency and openness are provided in Table 8 based on "The TOP Guidelines". (TOP, 2015; Nosek et al., 2015)

Table 6. List of relevant parameters for defining objectivity of a measurement

Dimension	Description
Risk of Bias	The degree to which the experimentation is designed and implemented to reduce the probability of bias. This includes elements of randomization, blind analysis, and operation/instrumentation bias (see Table 7 for additional details).
Transparency	The degree to which the experimental design, methods, and materials are made available for assessment (see Table 8 for additional details).

Table 7. Considerations regarding Evaluating Risk of Bias in a measurement

Parameter	Description
Selection Bias due to inadequate Randomisation	Systematic differences between baseline characteristics resulting from inadequate randomisation of measurements and conditions.
Performance Bias	Systematic differences between groups due to improper blinding of samples and the potential for preferential treatment by operators.
Detection Bias	Systematic differences between groups due to improper blinding of samples and the potential for bias in outcome identification by assessor.
Attrition Bias due to Incomplete Outcome Data	Systematic differences between groups due to incomplete data and imbalance amongst groups and/or outcomes.
Reporting Bias due to Selective Reporting	Systematic differences between reported and unreported findings based on pre-specified study/analysis plan.

Table 8. Considerations regarding Evaluating Transparency and Openness for a measurement

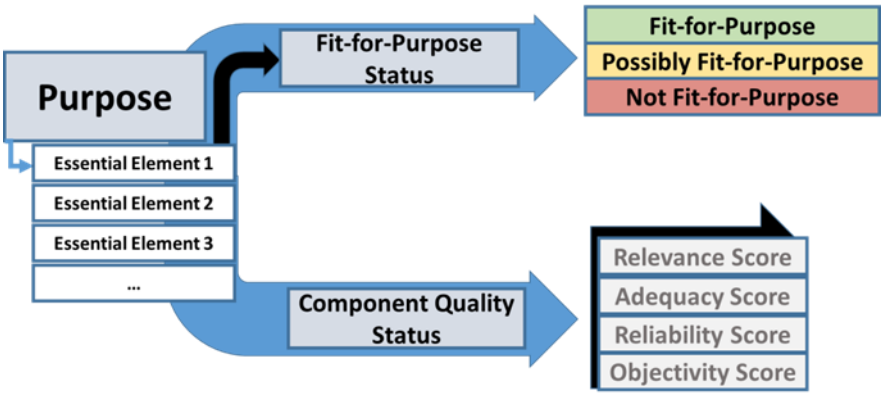
Parameter	Description
Data Transparency	The raw data and relevant metadata is made available for independent assessment.
Analytical/ Computational Method Transparency	The analytical method used to compute reported data applied in the final data analysis is made available for independent assessment.
Research Material Transparency	The materials applied in the study are made available for independent assessment.
Design and Analysis Transparency	The experimental design and analysis protocols are detailed and made available for independent assessment.
Pre-Registration of Study	The study plan is registered prior to being conducted and made available.
Pre-Registration of Analysis Plan	The data analysis plan is registered prior to being conducted and made available.

INTEGRATED ASSESSMENT AND SCORING PROTOCOL

21. The guiding principles are intended to be applied to achieve two goals: (1) to determine the 'fit-for-purpose' status for a given dataset or method, and (2) to provide a transparent and structured evaluation tool to facilitate ranking and selection of methods, protocols or data sources. This is also intended to assist in clarifying deficiencies and benefits of different method or data options with respect to the identified purpose.

22. A mixed tiered approach towards assessment for the evaluation of the guiding principle parameters is suggested to facilitate ease of use while allowing for a flexible yet sufficiently prescriptive methodology to enable quality ranking. The assessment of quality and fit-for-purpose attributes can be complicated and a range of approaches have been suggested in the literature. Quantitative methods such as multi-criteria data analysis (MCDA) are powerful but can be complicated to customize and implement for non-experts. On the other hand, qualitative tools tend to be easier to use but often lack sufficient resolution to provide guidance across a diverse array of scenarios (Martin et al. 2018). Herein, a two-step hybrid scoring assessment is proposed to simplify the assessment process while maintaining core elements of the MCDA process to enable ranking. The first step of the approach is to determine the fit-for-purpose classification of the method or data set. This process is solely dependent on determining if the minimum requirements set forth in the purpose 'are met', 'are possibly met' or 'are not met'. A second step is employed to determine a quantitative quality ranking across all of the parameters within the guiding principles for a given method or data set. The collective process provides the user with (1) an assessment of whether or not the needs of the purpose are met and (2) a generalized quality index that can also be applied to identify areas of improvement in the data reporting, method, and study design while further delineating differences between methods within or across categories. A schematic overview the process is depicted in Figure 2.

Figure 2. Overview of the Guiding Principle assessment and scoring protocol



Fit-for-Purpose Status

23. The status of a method or dataset as 'fit-for-purpose', 'possibly-fit-for-purpose' or 'not-fit-for-purpose' relies heavily on a well-described and sufficiently populated purpose statement reflecting the elements provided in Table 2. The purpose needs to be both well-defined and backed by sufficient information to enable the identification of 'fit-for-purpose' method(s) and dataset rather than only 'possibly-fit-for-purpose' outcomes. To further assist in identifying whether the purpose is sufficiently defined, a subset of commonly required parameters has been identified where there should be no ambiguity in the status of the method or data set. These parameters are provided in Table 8 along with pertinent questions that should be answerable with 'yes' or 'no'. If the information provided is insufficient for a decisive answer (which may simply indicate that the parameter is not applicable) then the *Purpose* needs to be better defined and consultation with the data requestor/user is recommended.

24. Once the purpose is suitably defined, the evaluation process is relatively simple. If all essential elements for the purpose are met by the method or dataset then it is classified as 'fit-for-purpose'. If one or more of the identified essential elements for the purpose are not met, then it is classified as 'not-fit-for-purpose'. For all other cases, there is not enough information yet to determine the method status and it is determined to be 'possibly-fit-for-purpose'. Note that the fit-for-purpose status only indicates that methodology or data set meets the request and does not explicitly comment on the quality of the method or dataset. The overall fit-for-purpose evaluation process is provided in Figure 3. In general, fit-for-purpose data sets are preferred, followed by possibly fit-for-purpose methods/data.

Table 9. Common Parameters that Should Sufficiently be Defined by the Measurement Purpose

Parameter	Description	Evaluation		
Substance Similarity	The measured substance represents the intended substance(s).	Yes	No	Maybe
Media & Environmental Conditions	The test media and environmental conditions are appropriate for the intended purpose.	Yes	No	Maybe
Sample Preparation	The sample preparation methodology is adequate for the intended purpose.	Yes	No	Maybe
CTQ Compliance	Are the identified CTQs are met by the data or method.	Yes	No	Maybe
Analytical Process Description	The experimental details are sufficient to allow complete reproduction of the experiment with an acceptable level of ambiguity. This includes sufficient descriptions of environmental conditions, applied media, sample preparation methodology, and measurement timescales.	Yes	No	Maybe
Equipment Consistency	The method is not known to suffer from significant manufacturer to manufacturer or instrument-to-instrument variability in the range of the resolution specified for the measurement, or steps have been taken to account for this variability through the use of reference standards, prescribed parametrisation, etc.	Yes	No	Maybe
Substance Description	The description of the substance origin (i.e. lot #, date of manufacture, synthesis route) is reasonably complete.	Yes	No	Maybe
Accuracy	The technique meets the specified accuracy or trueness requirement. If no requirement is specified, the measured value is within the limits of a reference material (if no reference material is available assume within 10% of an expected value). The accuracy of an analytical procedure expresses the closeness of agreement between the determined value and the true value or an accepted reference value.	Yes	No	Maybe
Resolution limit	The measurement system has the capability to detect and faithfully indicate small changes of the parameter within the resolution specified in the purpose.	Yes	No	Maybe
Robustness	The results of the method are not significantly impacted (effect less than specified resolution) by small, but definite variations in non-specified method parameters (e.g. user selectable data interpretation analysis options – chosen assumptions – and signal filtering methods). (An indicator of reliability under normal usage).	Yes	No	Maybe
Range	The range of the technique matches the required range. *The range of an analytical procedure is the interval between the upper and lower amounts of a parameter for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.*	Yes	No	Maybe

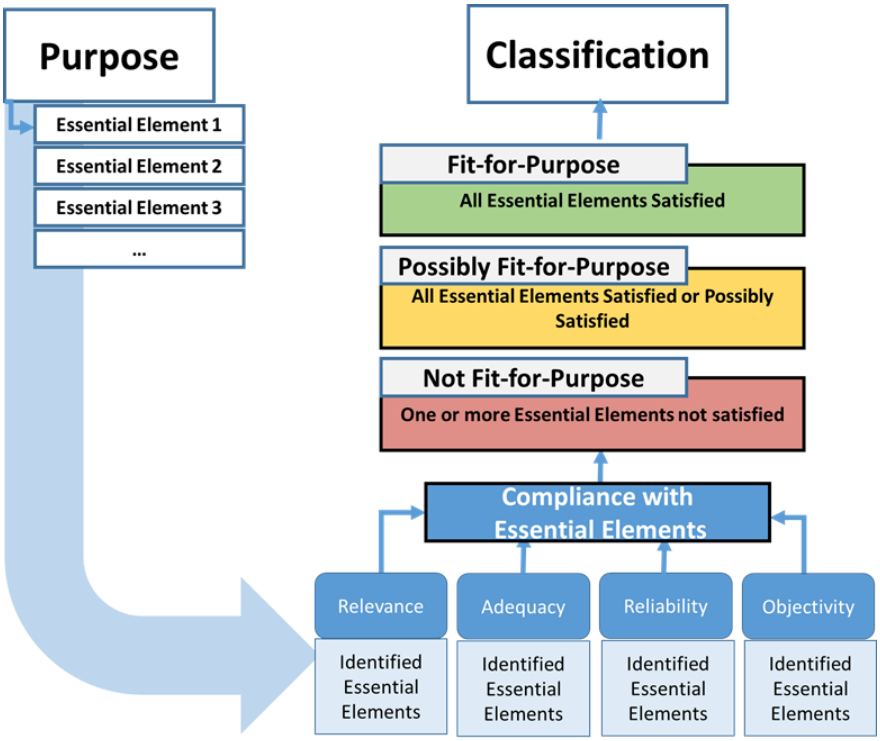


Figure 3. Overview of the Fit-For-Purpose Status Assessment Process

Component Quality Scores

25. In contrast to the fit-for-purpose status evaluation, the quality score includes contributions from all parameters in the evaluation components regardless of whether the parameter is "essential for purpose" (see Worksheets in Appendix 1). The component quality score is intended to provide perspective of the overall level of quality versus the ideal scenario.

26. For each guiding principle parameter that is suitably met a point is given, for each parameter that is not met a point is taken away. No points are given to parameters that remain uncertain. This is done for the full series of parameters across the guiding principles and results in component scores for each guiding principle (i.e. a component score for Relevance, Adequacy, Reliability, and Objectivity). To simplify this process a series of worksheets and instructions have been developed and are provided in Appendix 1. By answering questions related to questions in the worksheets (yes = parameter has been met; no = parameter has not been met) the component scores and subsequently the quality score can be derived. The quality score identification process through the use of these worksheets is identified in Figure 4.

27. For a given Fit-for-Purpose Status, the numerical quality component scores can help in further differentiating methods/data sets. By reviewing the component scores, strengths and areas for improvement can be identified as well as strategies to improve on the methodologies.

Ranking

28. For each classification, the method/data options can be ranked using the component scores depending on the overall purpose of the evaluation. The better the fit-for-purpose status and the higher the score the better suited the method.

29. The ranking philosophy for the methods/data sets is use specific and should be devised and documented prior to proceeding with the evaluation process to avoid bias. In many cases, the process of defining the purpose would also determine if one or more of the guiding principles are more important to fulfil the objectives. If so, the component scores for those guiding principles would have preference.

30. It is recommended that the evaluation proceeds by first ranking methods/data by the fit-for-purpose status first, then by the component quality scores. When a specific priority order of the component quality scores has not been established it is recommended that the evaluation proceeds by first ranking methods/data by the Fit-for-Purpose Status, then by the Relevance Component Score, the Adequacy Component Score and finally the Objectivity Component Score. To this effect, the following ranking format is suggested:

31. Fit-for-Purpose Status: (Relevance Score: Adequacy Score: Reliability Score: Objectivity Score)

32. By reviewing the parameters that were answered with a "maybe" or "no", adjustments may be feasible to improve the overall method. In general, the component scores should be used to guide process corrections for method/data improvement.

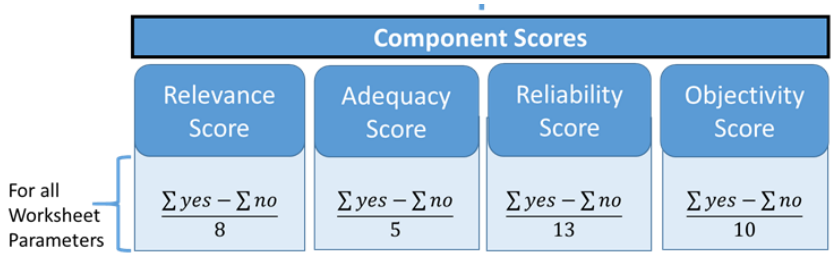


Figure 4. Schematic overview of the quality score and component score evaluation process for the worksheet found in Appendix 1.

33. As with any scoring system, the proposed methods require testing to validate or refine methodologies to promote consistency and ease of use. The above suggestions are simply a starting place for further discussion.

34. It is noted that the scoring philosophy is based on a simplified Multi-Criteria Decision Analysis (MCDA) structure. MCDA is a decision-making tool that divides decisions into smaller, more understandable parts, which are individually analysed then integrated to produce a meaningful solution. The decision making is broken down into essential elements that must be met, as well as general quality criteria. Rather than a fully numerical approach, a hybrid Boolean and numeric approach was chosen for simplicity (yes, no and maybe choices graded as by +1, -1, and 0, respectively; and the identification of essential elements to determine fit-for-purpose status and prioritising component quality scores instead of applying weighting factors). MCDA methods have been used successfully by Becker et al. (2017) and Linkov et al. (2011) for both similar (i.e., Nanosafety applications) and diverse applications. Linkov et al. (2007) applies

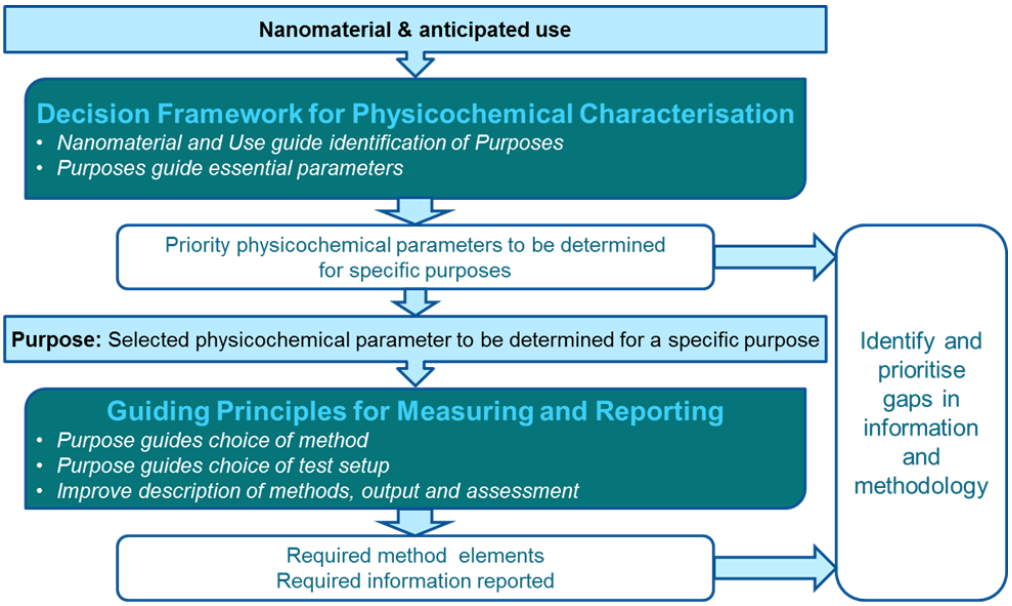
MCDA for Nanomaterial Environmental, Health and Safety Purposes. An advantage of the methodology is the ability to link performance information to decision criteria allowing for visualisation of the trade-offs involved in the decision-making process.

INTEGRATION WITH THE PHYSICO-CHEMICAL CHARACTERIZATION DECISION FRAMEWORK

35. The guiding principles are intended to serve as a generalised philosophy for ensuring that the purposes identified in the decision framework are specified to a suitable degree to facilitate appropriate method identification. They might also, for instance in the case of analogue data procurement, provide a means for identifying and highlighting the most relevant, adequate, reliable and objective data that meets the intended scope of the analogue criteria. A schematic illustration for the integration with the Physico-chemical Parameter Decision Framework is provided in **Figure 5**.

36. The proposed methodology and scoring regimen is anticipated to be refined through case studies developed from key purposes identified in the Decision Framework document. **By comparison of aggregate scores, the best available method for explicit purposes in the Decision Framework will be determined**. Additionally, in future editions of this document, recommendations will be provided to guide researchers to focus on meaningful method development and improved reporting practices to promote data usability and transparency. Facilitating the reproducibility of scoring amongst multiple scorers. It is proposed to include base templates for known methods to help facilitate agreement. The example base templates are found in Appendix 2.

Figure 5. Overview of input and output as well as interlinkages of the Physico-chemical Characterisation Decision Framework and the Guiding Principles for Measuring and Reporting



REFERENCES

1) Becker, R.A., Dellarco, V., Seed, J., Kronenberg, J.M., Meek, B., Foreman, J., Palermo, C., Kirman, C., Linkov, I., Schoeny, R. and Dourson, M., 2017. Quantitative weight of evidence to assess confidence in potential modes of action. *Regulatory Toxicology and Pharmacology*, *86*, pp.205-220.

2) Begley CG, Ellis LM: Drug development: raise standards for preclinical cancer research. Nature 2012,483(7391):531–533.

3) Beronius A, Molander L, Ruden C, Hanberg A: Facilitatin the use of non-standard in vivo studies in health risk assessment of chemicals: a proposal to improve evaluation criterial and reporting. Journal of Applied Toxicology. 2014; 34: 607-617.

4) Bioletta GS, Milner AM, Boyd IL: Quality assessment tools for evidence from environmental science. Environmental Evidence, 2014, 3:14, 1-14.

5) Bohannan J: Who's afraid of peer review? Science 2013, 342: 60–65.

6) Centre for Reviews and Dissemination: Systematic reviews: CRD's guidance for undertaking reviews on health care. 2009.

7) Collaboration for Environmental Evidence: Guidelines for systematic review and evidence synthesis in environmental management. Version 4.2. 2013.

8) ECHA: Chapter R.4: Evaluation of Available Information. Guidance on Information Requirements and Chemical Safety Assessment. Version 1.1, December 2011, 1-16. https://echa.europa.eu/documents/10162/13643/information_requirements_r4_en.pdf/d6395ad2-1596-4708-ba86-0136686d205e

9) Godlee F, Gale CR, Martyn CN: Effect on the quality of peer review of blinding reviewers and asking them to sign their reports: a randomized controlled trial. Jama 1998,280(3):237–240.

10) Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Sterne JA: The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. Br Med J 2011, 343: d5928.

11) International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use: Validation of Analytical Procedures: Text and Methodology Q2(R1). Step 4 Version, November 2005. 1-13 http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q2_R1/Step4/Q2_R1__Guideline.pdf

12) Linkov, I. and Moberg, E., 2011. *Multi-criteria decision analysis: environmental applications and case studies*. CRC Press.

13) Linkov, I., Bates, M.E., Trump, B.D., Seager, T.P., Chappell, M.A. and Keisler, J.M., 2013. For nanotechnology decisions, use decision analysis. *Nano Today*, *8*(1), pp.5-10.

14) Linkov, I., Satterstrom, F.K., Steevens, J., Ferguson, E. and Pleus, R.C., 2007. Multi-criteria decision analysis and environmental risk assessment for nanomaterials. *Journal of Nanoparticle Research*, *9*(4), pp.543-554.

15) Marchese Robinson RL, Lynch I, Peijnenburg W et al.: How should the completeness and quality of curated nanomaterial data be evaluated? Nanoscale, 2016, 8, 9919-9943.

16) Martin P, Bladier C, Meek B, et al. Weight of Evidence for Hazard Identification: A Critical Review of the Literature, Environ. Health Persp. 2018, 126(7):1-15. https://doi.org/10.1289/EHP3067

17) Nosek BA, Alter G, Borsboom D et al.: Promoting and open research culture, Science, 2015, 348, 1422-1425.

18) Oxman AD, Guyatt GH: The science of reviewing research. Ann N Y Acad Sci 1993, 703: 125–133.

19) Prinz F, Schlange T, Asadullah K: Believe it or not: how much can we rely on published data on potential drug targets? Nat Rev Drug Discov 2011,10(9):712–712.

20) Samuel GO, Hoffman S, Wright RA, Lalu MM et al.: Guidance on assessing the methodologic and reporting quality of toxicologically relevant studies: A scoping review. Environmental International. 92-93 (2016) 630-646.

21) TOP: The TOP Guidelines, Version 5, 2015,: http://cos.io/top

22) US EPA: Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency. Information Quality Guidelines. EPA/260R-02-008. October 2002. 1-56. https://www.epa.gov/sites/production/files/2017-03/documents/epa-info-quality-guidelines.pdf

23) US HHS, CDER, CBER: Analytical Procedures and Methods Validation for Drugs and Biologics - Guidance for Industry. Pharmaceutical Quality/CMC, July 2015. 1-13. https://www.fda.gov/downloads/drugs/guidances/ucm386366.pdf

24) Verhagen AP, de Vet HC, de Bie RA, Kessels AG, Boers M, Bouter LM, Knipschild PG: The Delphi list: a criteria list for quality assessment of randomized clinical trials for conducting systematic reviews developed by Delphi consensus. J Clin Epidemiol 1998,51(12):1235–1241.

25) Wells K, Littell JH: Study quality assessment in systematic reviews of research on intervention effects. Res Soc Work Pract 2009,19(1):52–62.

Appendix 1: Guiding Principle Worksheets and Evaluation Methods

Steps for Using the Worksheets

37. The use of the worksheets involves several steps:

    1) Complete the **Purpose** Worksheet ***(Sheet 1)***.
    2) Review the worksheets for **Relevance** ***(Sheet 2)***, **Adequacy** ***(Sheet 3)***, **Reliability** ***(Sheet 4)*** and **Objectivity** ***(Sheet 5)***.
    3) Check the "**Essential for Purpose**" box for line items that are specifically noted in the purpose worksheet ***(Sheet 1)***. Note that some boxes are pre-checked as commonly being essential for purpose.
    4) Identify if the purpose is sufficiently defined. If all pre-checked boxes can only be answered by "maybe" then the ambiguity in the purpose should be reduced through consultation with appropriate stakeholder (e.g. a representative of the party receiving the data).
    5) Fill out the worksheets, to evaluate a method or dataset.
    6) Identify the **Fit-for-Purpose Status and the Component Quality Scores** ***(Sheet 6)***.
    7) Rank the results and identify the best data sets and methods. Note, potential improvements to existing methodologies.

Step 1: Define the Purpose

38. The identification and conveyance of purpose lays the foundation for the evaluations. It cannot be understated that clearly defining the purpose and its bounds is critical. In terms of physico-chemical parameters, the purpose for a given measurement can vary widely. Sufficient detail and context, including both informative and quantitative parameters, is required to facilitate practical fit-for-purpose determination and reporting. Each question within each parameter should be answered as best as possible. The column identified as "worksheet parameter" provides an indication of the potentially impacted parameters from the other worksheets (R=Relevance, A=Adequacy, Ry=Reliability, and O=Objectivity).

Sheet 1. Purpose Worksheet.

Parameters	Questions defining purpose	Worksheet Parameter
Information Requirement	Why is the information required?	
Intended Use	How will the information be used? How will the information not be used?	
Substance Identification	For which substance(s) will the information apply? What parameters are used to distinguish similar and dissimilar substances of the same material class?	R.2, A.3
Relevant Realistic Scenarios	What realistic scenarios would the data aid in predicting, if any? What are the relevant conditions for that scenario? Is the information intended to be broadly or narrowly applied?	R.3, R.5, R.6, R.7, A.1, A.4
Critical to Quality (CTQ) Components	Are there specific key factors, implicated as critical to meeting the overall purpose? (e.g. method cost, availability requirements) 1. 2. 3.	R.8
Resolution Requirement	What is the required resolution of the data? (e.g. the practical quantification difference to be made against another substance – how small of a difference must be quantified). Precision Specification:	A.2, A.5, Ry.3, Ry.6
Trueness and Range Requirement	How accurate (true) is the data required to be and over what quantification range? Relevant Reference Standard: Identified Gold Standard Method: Trueness Specification: Required Range:	A.5, Ry.1, Ry.5

Step 2: Review Other Worksheets and Mark Line Items Defined in the Purpose Worksheet as "Essential to Purpose".

39. This step identifies essential parameters specified by the purpose for assessment for categorization purposes. Relevant line items/parameters in other worksheets are identified in the worksheet parameter column. See Sheet 2, Sheet 3, Sheet 4, Sheet 5.

Step 3: Identify if the purpose is sufficiently defined.

40. If all checked boxes can only be answered by "maybe" then the ambiguity in the purpose should be reduced through consultation with an appropriate stakeholder (e.g. a representative of the party receiving the data).

Step 4: Complete the Relevance, Adequacy, Reliability and Objectivity Worksheets.

Relevance

Sheet 2. Relevance Worksheet

Parameters		Essential for Purpose	Description	Evaluation		
R.1	Scientific Principle		The measurement directly measures the desired parameter and does not require broad assumptions or fitting parameters.	Yes	No	Maybe
R.2	Substance Similarity	√	The measured substance represents the intended substance(s)? *(see Purpose Worksheet/Substance Identification)*	Yes	No	Maybe
R.3	Media & Environmental Conditions	√	The test media and environmental conditions are appropriate for the intended purpose? *(see Purpose Worksheet/Relevant Realistic Scenario)*	Yes	No	Maybe
R.4	Sample Preparation	√	The sample preparation methodology is adequate for the intended purpose? *(see Purpose Worksheet/Relevant Realistic Scenario)*	Yes	No	Maybe
R.5	Sample Concentration		The concentration of sample during measurement is sufficient and adequate for the intended purpose? *(see Purpose Worksheet/Relevant Realistic Scenario)*	Yes	No	Maybe
R.6	Time		The timescale for conducting the experiment is appropriate for the intended purpose? *(see Purpose Worksheet/Relevant Realistic Scenario)*	Yes	No	Maybe
R.7	History		The sample storage and handling/processing prior to measurement appropriate are appropriate.	Yes	No	Maybe
R.8	CTQ Compliance	√	Are the identified CTQs met by the data or method?	Yes	No	Maybe
			Total answers for "Essential for Purpose" parameters			
			Total answers for all parameters			
			Relevance Component Score ((Total "Yes" – Total "No")/8)	RV=_____		

Adequacy

Sheet 3. Adequacy Worksheet

Parameters		Essential for Purpose	Description	Evaluation		
A.1	Analytical Process Description	√	The experimental details are sufficient to allow complete reproduction of the experiment with an acceptable level of ambiguity. This includes sufficient descriptions of environmental conditions, applied media, sample preparation methodology, and measurement timescales.	Yes	No	Maybe
A.2	Equipment Consistency	√	The method is not known to suffer from significant manufacturer to manufacturer or instrument-to-instrument variability in the range of the resolution specified for the measurement, or steps have been taken to account for this variability through the use of reference standards, prescribed parametrization, etc.	Yes	No	Maybe
A.3	Substance Description	√	The description of the substance origin (i.e., lot #, date of manufacture, synthesis route) is reasonably complete.	Yes	No	Maybe
A.4	Storage and Handling Description		The description of the storage and handling of the substance after initial receipt and prior to analysis is thorough.	Yes	No	Maybe
A.5	Uncertainty Budget		A critical assessment of the source of experimental errors and uncertainty have been contemplated and conveyed for the method indicating suitability.	Yes	No	Maybe
			Total answers for "Essential for Purpose" parameters			
			Total answers for all parameters			
			Relevance Component Score ((Total "Yes" – Total "No")/5)	A=_____		

Reliability

Sheet 4. Reliability Worksheet

Parameters		Essential for Purpose	Description	Evaluation		
Technique Related:						
Ry.1	Accuracy	√	The technique meets the specified accuracy or trueness requirement. If no requirement is specified, is the measured value within the limits of a reference material? If no reference material is available assume within 10% of an expected value. The accuracy of an analytical procedure expresses the closeness of agreement between the determined value and the true value or an accepted reference value.	Yes	No	Maybe
Ry.2	Specificity		The method measures the desired parameter unequivocally, and is not influenced by confounding factors under the measurement conditions.	Yes	No	Maybe
Ry.3	Resolution limit	√	The measurement system has the capability to detect and faithfully indicate small changes of the parameter within the resolution specified in the purpose.	Yes	No	Maybe
Ry.4	Linearity		The method produces a result directly proportional to the quantity of a parameter over the range of the technique. This also includes goodness of correlation of non-linear functions.	Yes	No	Maybe
Ry.5	Range	√	The range of the technique matches the required range. *The range of an analytical procedure is the interval between the upper and lower amounts of a parameter for which it has been demonstrated that the analytical procedure has a suitable level of precision, accuracy and linearity.*	Yes	No	Maybe
Ry.6	Robustness	√	The results of the method are not significantly impacted (effect less than specified resolution) by small, but definite variations in non-specified method parameters (e.g. user selectable data interpretation analysis options – chosen assumptions – and signal filtering methods). (An indicator of reliability under normal usage).	Yes	No	Maybe
Laboratory Related:						
Ry.7	Operator Training		The laboratory applies formal validation protocols for the qualification of operator performance.	Yes	No	Maybe

Parameters		Essential for Purpose	Description	Evaluation		
Ry.8	Instrument Validation		The laboratory implements regular instrument calibrations and controls to identify deviations in performance.	Yes	No	Maybe
Ry.9	Media Description / Characterization		The test media applied in analysis (if any) has been characterized or described to facilitate independent comparative measurements.	Yes	No	Maybe
Ry.10	Reference Measurements		Reference material measurements have been applied to aid in normalizing results or for confirming operability.	Yes	No	Maybe
Ry.11	Intermediate precision		Within laboratory variations in reported outcomes amongst available equipment and operators are reported.	Yes	No	Maybe
Ry.12	Between Laboratory Precision		The laboratory practices for the parameter have been confirmed via comparison with results from external laboratories (e.g. participation in inter-laboratory comparisons, independent certification, etc.).	Yes	No	Maybe
Ry.13	Experience Level		The laboratory staff routinely performs the technique and has expert or intermediate-level understanding of its fundamental principles.	Yes	No	Maybe
			Total answers for "Essential for Purpose" parameters			
			Total answers for all parameters			
			Relevance Component Score ((Total "Yes" – Total "No")/13)	Ry=_____		

Objectivity

Sheet 5. Objectivity Worksheet

	Parameter	Essential for Purpose	Description	Evaluation		
	Risk of Bias					
O.1	Selection Bias due to inadequate Randomization		Systematic errors were avoided through randomization of samples. Sample order was randomized?	Yes	No	Maybe
O.2	Performance and Detection Bias		Samples were blinded? Systematic differences between measurements due to improper blinding of samples and the potential for preferential treatment by operators was avoided.	Yes	No	Maybe
O.3	Attrition Bias due to Incomplete Outcome Data		All data sets used were complete? Significant measurement information differences between samples were avoided.	Yes	No	Maybe
O.4	Reporting Bias due to Selective Reporting		Results were not selected based on outcome and pre-specified selective reporting is not evident in study/analysis plan. Were all findings were reported?	Yes	No	Maybe
	Transparency					
O.5	Data Transparency		The raw data and relevant metadata is made available for independent assessment.	Yes	No	Maybe
O.6	Analytical Code Transparency		The analytical code applied for data analysis is made available for independent assessment.	Yes	No	Maybe
O.7	Research Material Transparency		The materials applied in the study are made available for independent assessment.	Yes	No	Maybe
O.8	Design and Analysis Transparency	√	The experimental design and analysis protocols are detailed and made available for independent assessment	Yes	No	Maybe
O.9	Pre-Registration of Study		The study plan is registered (documented) prior to being conducted and made available.	Yes	No	Maybe
O.10	Pre-Registration of Analysis Plan		The data analysis plan is registered (documented) prior to being conducted and made available.	Yes	No	Maybe
			Total answers for "Essential for Purpose" parameters			
			Total answers for all parameters			
			Relevance Component Score ((Total "Yes" – Total "No")/10)	O=_____		

Step 5: Identify the Classification and the Quality Score

Determine the classification of the method by evaluating the outcomes of "essential to purpose" parameters across the four worksheets (i.e. Sheet 2, Sheet 3, Sheet 4, Sheet 5). If all parameters are answered with a "Yes" then the method or data set is classified as "fit-for-purpose". If one or more parameters is answered with a "No" then the method or data set is classified as "Not fit-for-purpose". Else, the method is classified as "possibly fit for purpose".

41. To determine the quality score, the component scores need to be added. Sheet 6 below can be filled out to derive the quality score by summing the component scores.

Sheet 6. Quality Score

Guiding Principle	**Quality Score**
Relevance	Relevance Component Score_____
Adequacy	Adequacy Component Score_____
Reliability	Reliability Component Score_____
Objectivity	Objectivity Component Score_____

1. For each classification, the method/data options can be ranked using the component scores. Priority and component importance (e.g., Reliability versus Objectivity) depends on the purpose for the evaluation. The default scenario has the order of importance as Relevance, Adequacy, Reliability followed by Objectivity. The better the fit-for-purpose status and the higher the component scores (in order of priority) the better suited the method.

42. The ranking philosophy applied for the methods/data sets should be devised and documented prior to proceeding with the evaluation process to avoid bias. In many cases, the process of defining the purpose would also determine if one or more of the guiding principles are more important to fulfil the objectives. If so, then the Component Scores for those guiding principles would gain priority.

43. When a specific priority order is not identified, it is recommended that the evaluation proceeds by first ranking methods/data by the Fit-for-Purpose Status, then by the Relevance Component Score, the Adequacy Component Score and finally the Objectivity Component Score. To this effect, the following ranking format is suggested:

    Fit-for-Purpose Status: Quality Score (Relevance Score: Adequacy Score: Reliability Score: Objectivity Score)

44. By reviewing the parameters that were answered with a "Maybe" or "No", adjustments may be feasible to improve the overall method. In general, the Quality Score and Component Scores should be used to guide process corrections for method/data improvement.

Appendix 2: Base Templates

1. Base templates are common information sets that pertain to specific elements within an evaluation scheme. For standard methodologies and analyses that utilise commercial analytical equipment, elements of the Guiding Principles can be pre-populated in templates that should assist in the selection of appropriate methods once the purpose is defined. A similar practice can be conducted for laboratory assessments based on normal practices and established protocols.

45. An example analytical technique base template is given below. A similar template can be used to cover common laboratory procedures for data handling and analysis.

Parameters	Description
TECHNIQUE RELATED:	
• **Measured Parameter**	e.g. hydrodynamic diameter
• **Physical Principle used to obtain parameter**	e.g. Stoke's sedimentation
• **Alternative Uses**	e.g. Sedimentation Rate Calculations if applied density is known or via raw data
• **Reported Accuracy**	e.g. +/- 5 nm
• **Reference Materials Applied for Accuracy Claim**	e.g. NIST SRM XXXX
• **Reported Resolution limit**	e.g. 5 nm
• **(Reference) Materials Applied for Resolution Claim**	e.g. ~ monodisperse iron oxide dispersion measured by TEM
• **Reported Intermediate Precision**	e.g. 3 nm
• **(Reference) Materials Applied for Precision Claim**	e.g. NIST SRM XXXX
• **Linearity**	• e.g. 99.5% least squares residual to linear fit
• **(Reference) Materials Applied for Linearity Claim**	e.g. ~ monodisperse iron oxide dispersion measured by TEM
• **Range**	e.g. ~5 to 10000 nm
• **Robustness**	e.g. Requires knowledge of adjusting for baseline detector drift and impacted by inappropriate warm up routines, sample cell cleaning
• **Equipment Cost**	e.g. ~$30,000 (estimated in 2015)
MATERIAL REQUIREMENTS:	
• **Major determining factor for method validity**	e.g. Must adsorb radiation at XXX nm; Samples must be concentrated above ~1% wt. Nanomaterials must be measured in liquid

Parameters	Description
• **Complexity of Method**	e.g. Requires moderate to expert laboratory operators
• **Nanomaterial and Fluid Compatibility Requirements**	• e.g. Some equipment allows for non-aqueous fluids. Nanomaterial must adsorb radiation at XXX nm.
• **Interlaboratory precision of Method in documented Round Robin**	e.g. ~10 nm as reported in XXX
• **Intermediate precision from Round Robin**	e.g. ~4 nm as reported in XXX