Guidance on the GRACIOUS Framework

for Grouping and Read-Across of Nanomaterials and Nanoforms

gracious

Grouping, Read-Across, CharacterIsation and classificatiOn framework for regUlatory risk assessment of manufactured nanomaterials and Safer design of nano-enabled products

Annex: GRACIOUS Framework Guidance Document

Contents

1.0 Introduction

The unique or enhanced properties of nanomaterials have led to their increased use in products across many sectors. However, these same properties have triggered concerns of enhanced hazard and risk (EUON). The variability of physicochemical parameters used to describe nanomaterials mean that a single substance may have many different nanoforms (NFs) (European Parliament, 2006) that in turn may display diverging fate, toxicokinetic, toxicological and ecotoxicological properties. Comprehensive testing of all of these properties for all nanoforms may make developing nanotechnology economically impractical. Grouping has been used to generate data, without the need to commission expensive testing, for chemicals lacking the information needed for risk assessment.

TIP for New Users

A nanoform is defined in Annex VI of REACH as "a nanoform is a form of a natural or manufactured substance containing particles, in an unbound state or as an aggregate or as an agglomerate and where, for 50 % or more of the particles in the number size distribution, one or more external dimensions is in the size range 1 nm-100 nm, including also by derogation fullerenes, graphene flakes and single wall carbon nanotubes with one or more external dimensions below 1 nm.

For this purpose, 'particle' means a minute piece of matter with defined physical boundaries; 'agglomerate' means a collection of weakly bound particles or aggregates where the resulting external surface area is similar to the sum of the surface areas of the individual components and 'aggregate' means a particle comprising of strongly bound or fused particles."

In other words, a nanomaterial such as TiO_2 may exist in different nanoforms that vary in size, shape or coating composition.

The project GRACIOUS (Grouping, Read-Across, CharacterIsation and classificatiOn framework for regUlatory risk assessment of manufactured nanomaterials and Safer design of nano-enabled products) was set up to develop and promote grouping and read-across approaches for nanomaterials. This guidance document will walk the user through the steps in the Framework (REF), demonstrating the

requirements of each step and how a user can make grouping and read-across decisions. Worked examples show how the approaches recommended in the Framework can be applied in real life.

1.1 Regulatory drivers for Grouping

Grouping, categorisation and read-across have been recommended by regulators as approaches to generate hazard endpoint data for substances lacking data, without needing to commission animal testing (Article 13 of REACH). Grouping allows similar chemicals to be placed within a group, from which the information from data rich members can be used to predict the hazard properties of members which lack data, using either read-across or by development of predictive *in silico* models. A group is therefore developed to enable identification and provision of relevant data for a specific endpoint. The principles of grouping require that:

- It is based on a sound scientific hypothesis.
- The hypothesis needs to link physicochemical properties and hazard.
- Scientific justification of the link is essential, which can be facilitated if there is knowledge of the Mechanism of Action (MoA) or Adverse Outcome Pathway (AOP).

Amendments to the Annexes of REACH were published in December 2018 and enforced from January 2020, requiring additional information to be included in the registration dossiers of substances which have one or several NFs that are placed on the market in the EU (European Parliament, 2006). The amendments require the identification and assessment of different NFs of the same substance if the total annual tonnage of the substance placed on the market by the manufacturer or importer exceeds 1 tonne and according to the requirements of the total tonnage level, regardless of the amount of the individual nanoform produced. ECHA has recommended the use of grouping to avoid the need for extensive animal testing of NFs that might only be placed on the market in low quantities, a basic principle of REACH. The GRACIOUS project is in part intended to investigate ways to achieve this goal.

1.2 Existing Grouping methods for nanoforms

There have been exploratory investigations into the use of grouping for NFs and the scientific justification required to validate grouping by organisations and projects such as:

- US National Institute on Occupational Safety and Health (NIOSH)
- US-Canada Regulatory Cooperation Council (RCC)
- ITS-NANO (EU FP7 project)
- MARINA (EU FP7 project)
- DF4NanoGrouping (ECETOC project)
- NANoREG and NanoReg2 (EU FP7 project)
- ECETOC NanoApp (specific to ‘Sets of nanoforms’)
- NanoGravur (German project)

The GRACIOUS Framework integrates the principles of this earlier work to produce a comprehensive structure for a user to be able to address their own grouping requirements.

1.3 Do you want to know more?

The following resources can provide more information:

European Chemicals Agency (2019). Appendix for nanoforms applicable to the Guidance on Registration and Substance Identification.

Available at https://echa.europa.eu/documents/10162/17250/how_to_register_nano_en.pdf/f8c046ec-f60b-4349-492b-e915fd9e3ca0

> *ECHA has published a number of guidance documents to support registrants of nanoforms under REACH. This document explains these registration obligations, how to distinguish nanoforms and the physicochemical characterisation required in a REACH registration dossier for a nanoform.*

European Parliament (2006). REGULATION (EC) No 1907/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH).

> *This is the legal text of the REACH regulation that has specific regulatory obligations for the manufacturers, importers and users of some nanomaterials. It also gives a legally binding definition of a nanoform that is applied throughout the GRACIOUS Framework.*

European Union Observatory for Nanomaterials (EUON). Nanomaterials and Health. Available at https://chemicalsinourlife.echa.europa.eu/nanomaterials-and-health?utm_source=euon.echa.europa.eu&utm_medium=display&utm_campaign=ux&utm_content=homepage-links

> *This short article introduces how nanoforms can reach parts of the body bulk forms cannot and why they need to be assessed separately under European legislation*

Stone, V., Gottardo, S., Bleeker, E., Braakhuis, H., Dekkers, S., Fernandes, T., Haase, A., Hunt, N., Hristozov, D., Jantunen, P., Jeliazkova, N., Johnston, H., Lamon, L., Murphy, F., Rasmussen, K., Rauscher, H., Jiménez, A. S., Svendsen, C., Spurgeon, D., Omen, A. G. (2020). A framework for grouping and read-across of nanomaterials- supporting innovation and risk assessment. Nano Today, 35, [100941]. doi.org/10.1016/j.nantod.2020.100941

> *This paper gives an introduction to the GRACIOUS Framework*

2.0 The GRACIOUS Framework

2.1 Aims of the Framework

While grouping is an established method for filling data gaps for the hazard and risk assessment of substances, its use for nanomaterials is less well established. The GRACIOUS Framework builds upon and combines concepts from multiple previous and current projects (section 1.2), as well as guidance from regulators. It is intended to:

- Support practical and evidence-based grouping of NFs by facilitating data gathering for hazard and risk assessment, risk management and related decision making, thereby meeting the needs of various global stakeholders, particularly regulators and industry.
- Develop a number of robust scientific arguments (so called pre-defined hypotheses) that justify grouping and read-across of NFs.
- Facilitate the development of new (so called user-defined) hypotheses to support grouping and read-across.
- Consider not only intrinsic physicochemical properties and (eco)toxicological effects, but also extrinsic (system-dependent) descriptors of exposure, toxicokinetic and environmental fate.
- Provide guidance on how the outputs can subsequently be used, and aligned to the initial purpose of grouping.
- Support decision making spanning regulatory risk assessment and safe innovation/Safe(r)-by-Design (SbD) of nano-enabled products.
- Apply the 3Rs principles in order to reduce, refine and replace animal testing for human health and environmental hazard assessment where possible, by supporting the use of grouping, read-across, modelling and *in vitro* testing.

2.2 Background to the GRACIOUS Framework

The GRACIOUS Framework was designed to integrate industrial (e.g. DF4NanoGrouping) and regulatory (e.g. ECHA) grouping concepts. In order to ensure that the Framework would be fit for purpose, the opinions of diverse stakeholder groups (spanning academia, regulation, industry, standardisation and NGOs among others) were sought and incorporated into the detailed design of the Framework following two rounds of stakeholder consultations involving collection of feedback via surveys and in-depth interviews. In the future, insights from the Framework will be incorporated by OECD into an updated edition of their Guidance Document on Grouping of Chemicals.

Tips for a new user

Interested to know more about our stakeholder engagement activities? Here you can find records of our completed open consultation activities: https://www.h2020gracious.eu/about/stakeholders

Brief video overviews of the Framework and its elements are also available at https://www.h2020gracious.eu/library/dissemination-materials

2.3 The GRACIOUS Framework Structure

A hypothesis-driven approach to grouping is essential in order to align with European legislation (e.g. REACH), but also to provide a scientific basis for any grouping decision. The Framework uses a stepwise approach, from which users can exit if they believe the grouping hypothesis has been accepted (or rejected) with the data at hand. When moving through the steps of the Framework, an increasing amount and complexity of data is required. The structure of the Framework requires only the data needed to support the grouping hypothesis and therefore scientifically justify grouping, thus avoiding unnecessary testing.

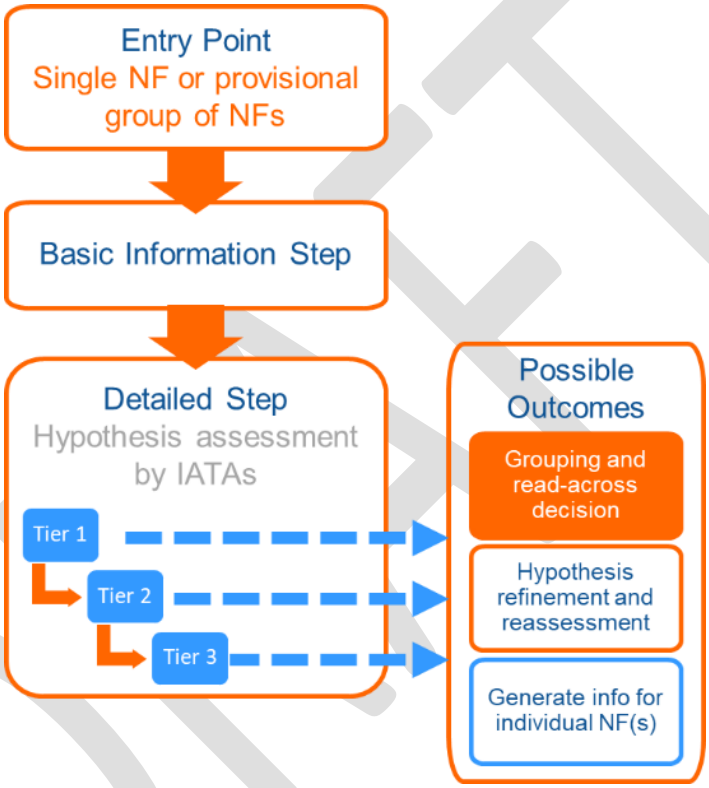


Figure 2.3.1: A simple overview of the GRACIOUS Framework demonstrating the Basic Information Step which provides an initial output of possible grouping hypotheses, followed by the more Detailed Step which gathers the evidence needed to accept or reject the proposed grouping hypotheses.

This simple description of the Framework demonstrates the importance of the user being very clear regarding which NF they are considering for assessment, and that a suite of basic information will be required for each NF before it is possible to continue further through the Framework.

The GRACIOUS Framework has defined a set of pre-defined grouping hypotheses, developed IATAs suitable for use with these hypotheses and executed case studies with NFs to test the utility of the Framework. It is recommended that the user first considers whether these pre-defined hypotheses apply to their NFs to reduce the time and resources required, but, if necessary, the user can define their own grouping hypothesis and associated IATA (Figure 2.3.2). The IATAs guide collation of the required information into a data matrix in order to support assessment of the similarity of the candidate NFs,

allowing the group members to be confirmed. The Framework allows for a variety of conclusions to be drawn, both quantitative and qualitative.

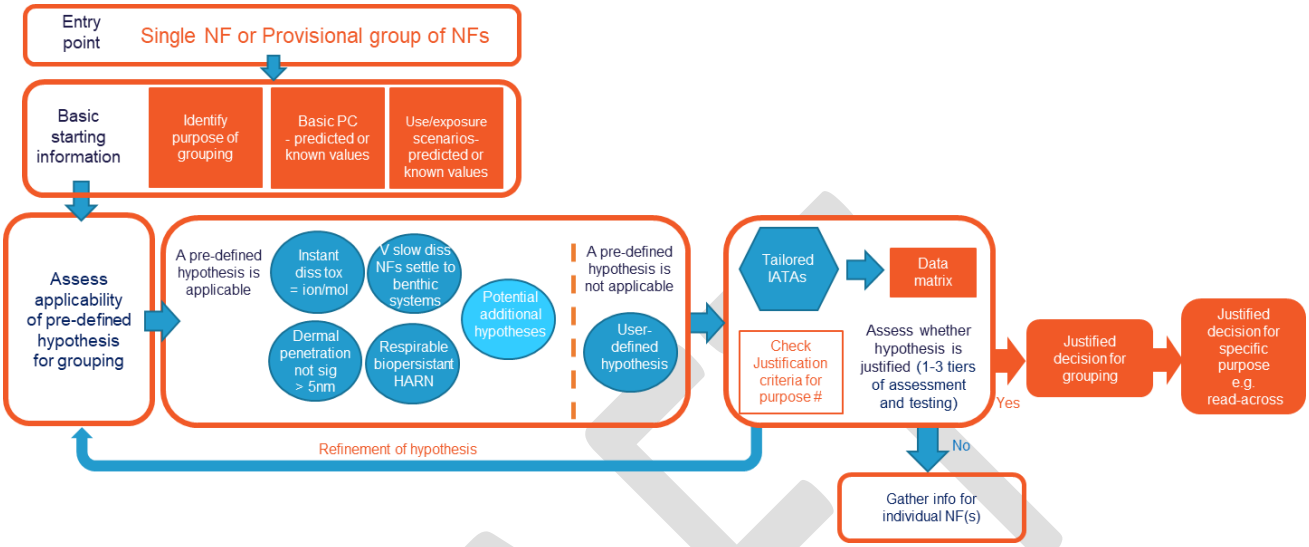


Figure 2.3.2: Detailed overview of the GRACIOUS Framework showing how pre-defined or user-defined hypotheses can be used to group candidate NFs.

This guidance document will walk the user through the steps in the Framework, demonstrating the purposes of each step and how the user can generate and assess the data applicable at each step to either move on through the Framework or to exit it where appropriate. This will be supported by brief worked examples undertaken by project partners and links to more detailed explanations for those interested.

Tips for a new user

What the Framework can…and cannot do

The GRACIOUS Framework can provide a structure for a user interested in grouping NFs to reach a scientifically justified conclusion in a systematic and logical way. It gives the user the tools and approaches needed to justify grouping and gives guidance on choosing the best one for their situation. A user can use the worked examples and case studies that have demonstrated how the Framework has been used to group NFs for different purposes in specific scenarios to support their approach. The Framework has been designed to integrate with other tools designed to aid those researching nanomaterials. However, a user must be aware that every grouping scenario will be unique so the Framework cannot give step-by-step instructions for every potential purpose and grouping hypothesis. Although some pre-defined hypotheses and their associated IATAs have been designed and tested, a user will need to use their own expertise to interpret their experimental results and how they relate to the purpose of the grouping exercise.

2.4 Do you want to know more?

The following resources can provide more information:

ECHA (2008). Guidance on information requirements and chemical safety assessment Chapter R.6: QSARs and grouping of chemicals. Available at https://echa.europa.eu/documents/10162/17224/information_requirements_r6_en.pdf/77f49f81-b76d-40ab-8513-4f3a533b6ac9

> *This guidance document introduces how grouping can be used to reduce the amount of animal testing to generate the data needed to satisfy REACH obligations for all chemicals.*

OECD (2017). Guidance on Grouping of Chemicals, Second Edition, OECD Series on Testing and Assessment, No. 194, OECD Publishing, Paris, https://doi.org/10.1787/9789264274679-en

> *This guidance document also introduces different grouping approaches for all chemicals and how data gaps can be filled using this approach. It explains how different types of chemical can be addressed by grouping including an initial consideration of grouping of nanomaterials.*

3.0 Using the GRACIOUS Framework

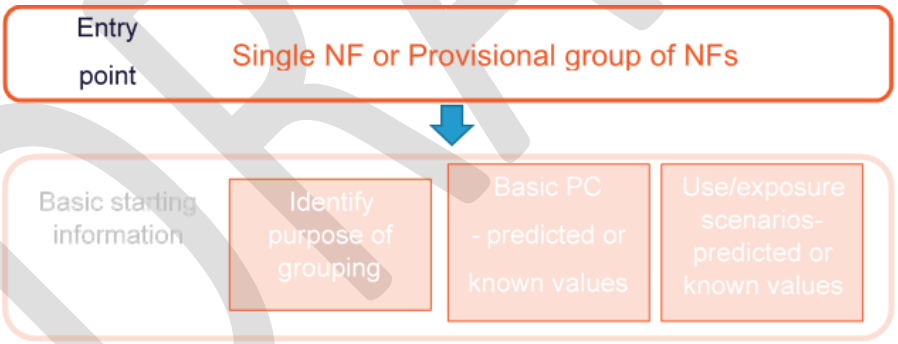
The GRACIOUS Framework can be applied by following the instructions below. Alternatively, GRACIOUS has generated a Blueprint of software that can be used to support grouping and read-across. This Blueprint is available as an open access PDF document and can be applied by software developers to implement the GRACIOUS Framework (or relevant parts) via risk assessment software tools (see section 3.5).

Tip for new users

Using estimated data in the GRACIOUS Framework:

Within the GRACIOUS Framework, the pre-defined hypotheses have been formulated to clearly indicate in the first half of the hypothesis sentence, the key physicochemical characteristics and exposure route. Basic information should be available for all candidate NFs, including target and (potential) source materials. Source materials are those NFs with existing hazard data. As this is likely to be historic data, possibly measured before analytical methods had been validated and standardised, these source materials might lack the PC characteristics specified for the basic information. In this situation, we recommend using estimated values which can be replaced with measured values at a later point in the process.

3.1 Entering the Framework: Which NF should be used?

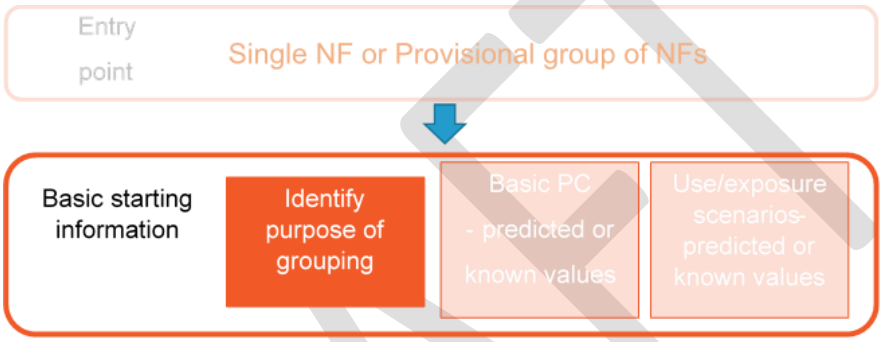


Firstly, the user needs to identify which NFs they are interested in. This is most easily done by identifying all the relevant NFs in their product portfolio, although it should be realized that these will not necessarily end up all in the same group for hazard/risk characterisation. If the user knows that there is little experimental data available for their candidate NF(s), especially if they hope to use read-across to fill in any data gaps, they could consider identifying a data-rich NF (or in some situations a non-nanoform) from outside their product portfolio. For a formal read-across for regulatory purposes, the NF(s) or non-NFs with data (usually animal data) suitable to support risk assessment would be referred to as the *source*, while the NF(s) lacking data would be the *target(s)*). The source(s) can be identified at different points of the Framework and these will be highlighted in this guidance.

3.2 The Basic Information Step

Grouping always needs a strong scientific justification, so a basic understanding of the NFs under assessment is essential before the next steps in the Framework can be undertaken. Many of the questions raised when progressing through an IATA will require that the physicochemical identity and lifecycle of each NF under consideration are well-understood. If these aspects are missing or poorly defined, drawing a scientifically justified conclusion on grouping will be impossible.

3.2.1 Purpose of Grouping



The Framework is designed to support at least three different potential purposes: Precautionary risk management, SbD approaches and Regulatory risk assessment.

'Precautionary risk management measures' and 'Safe(r) by design' approaches

The Framework can help to support both SbD approaches and precautionary risk management measures. The Framework allows the user to identify information from similar materials, used in similar applications, to improve the safety of new NFs and nano-enabled products, or to develop informed precautionary risk management measures.

SbD approaches can apply at any stage of the development process, for example, to reduce hazard(s) (e.g. by elimination of a hazardous substance and/or substitution for another one that is known to be safer). SbD approaches can also provide the information needed to reduce the release of nanomaterials from a product during its use, or from a process (e.g. coating of a NF to reduce dustiness). The application of such approaches allows incorporation of safety considerations early in the innovation of new materials or products, helping to make innovation more cost-effective.

Similarly, if the scientific knowledge and data are insufficient to assess the risk(s) of a NF in an exposure scenario, then precautionary risk management measures can be applied. Such measures aim to prevent or reduce exposure (and therefore risk) by implementing a conservative risk management plan, including measures such as engineering or administrative controls, or use of personal protective equipment.

The difference between the two purposes is relevant at the point of application of grouping, where the SbD approach is applied at the design phase, while the precautionary measures approach is likely to be

applied in the manufacturing and downstream use of a material/product that already exists but no REACH registration is required (regulatory requirements would always take precedence when applicable). Both approaches are for non-regulatory purposes, to allow users to make informed decisions to help avoid risks in a specific scenario.

For both purposes, during progression through the IATAs and the associated tiered testing strategies (Section 3.3), tier 1 tests are likely to be sufficient. For SbD approaches, tier 1 tests may also be sufficient for narrow group ranges (highly similar NFs) at the stages of REACH registration and launch. However, for users requiring more information and/or wider group ranges and/or higher certainty, incorporation of higher tiers of testing can be utilised.

For SbD approaches and identification of precautionary measures, data such as physicochemical characteristics can be estimated at the early innovation stages. For example, in the design phase, before production of a prototype, the designer might use grouping to identify lower risk options for development. Estimated values however would only be used at early stages of SbD, and at every Gate of the *Stage-Gate* process (Cooper, 2017) more information with greater certainty, about e.g. safety, needs to be provided. At the later stages of the SbD process, the information produced is likely to be usable for regulatory purposes.

Regulatory risk assessment approaches

EU chemical regulations, such as REACH (European Parliament, 2006), require information on specific hazard endpoints to be provided for a substance or mixture placed on the market. If information on these endpoints are not provided by an appropriate study performed on that substance, scientific justification is needed to allow the use of data from a different *source* (e.g. read-across from a different substance or NF). Under REACH, this applies to each NF of a substance placed on the market. Grouping is a scientifically justified method by which similar NFs can be identified, in order to allow read-across of data from group members possessing the information needed for risk assessment, to group members which lack this information. Application of grouping and read-across reduces the need for new animal studies.

The user can initially start the grouping process for SbD purposes in the early stages of developing a material/product and then progress to regulatory risk assessment in the later stages of the innovation process, prior to releasing the material/product to the market. However, grouping can only be applied in a regulatory context after changing any estimated values to measured/modelled/calculated ones. In fact, it is likely that this will be the natural course of events for many NFs.

For regulatory purposes, tier 1 tests are suitable where there is a high degree of similarity (and certainty) between all candidate members of the group. Note that tier 1 does not always mean lower confidence; it depends upon how similar the NFs are and the methods used for data acquisition. Tier 2 or tier 3 data can be used if tier 1 data are not available, lack quality, are too variable, have a lower degree of similarity, do not allow conclusions to be drawn, or when the user seeks wider group ranges. Expert judgement will be needed to ensure all of the evidence required for a regulatory purpose is available and included.

3.2.1.1 Instructions to define the purpose for grouping

Defining the purpose for grouping has a significant impact on how the user uses the Framework, so it is important that the purpose is well-considered. We recommend that the following steps are used.

1. Identify the general purpose of grouping

Within the GRACIOUS Framework, the user is prompted to choose one of several different purposes:

- To fill a data gap in a regulatory dossier.
- To develop precautionary risk management measures.
- To steer SbD innovation.

These purposes determine the level of detail and the type of information generated by the Framework so that the outputs are tailored to the user's purpose for grouping. An example of this would be whether to use estimated data in the IATA or not. If the purpose for grouping is SbD or to identify precautionary measures, the burden of proof may not be too high, so estimated data could suffice, especially if resources (time, cost) are limited during the development phase. If the purpose of grouping is to meet regulatory requirements, measured data is always preferred and estimated data is almost never acceptable in isolation. Modelled or calculated data can be used for some endpoints if their use is well justified, using validated models; such data will generally be part of a weight-of-evidence approach. The general purpose of the grouping will impact on the decision-making process through the Framework, such as whether higher tier studies are needed within a DN, or the degree of confidence needed for a similarity decision that defines the final group.

2. Elaborate the specific purposes of grouping relevant to the user

The user needs to provide context for the grouping to be conducted that help to explain their specific situation, for example:

Regulatory purposes

- Is the user a regulator or an industrial operator working to meet regulatory requirements?
- Is the intention to meet a single study requirement, a regulatory endpoint or all endpoints within a single route of exposure?

Safe(r)-by-Design purposes

- Can maintaining technical effectiveness of the product be balanced with safety by using grouping to inform hazard considerations?

Precautionary risk management measures

- How many locations are covered by the assessment?
- What is the level of training of operators?
- What are the operating characteristics of individual locations?

Note that the pre-defined hypotheses developed for the GRACIOUS Framework are intended to be general and not substance-specific, in order to have wide applicability. This requires that when the user undertakes grouping, they understand how their own specific context fits into these general structures. These factors could have an impact on the final grouping decision.

3. **Identify potential impacts of the purpose on the decision-making process through the Framework**

Although every user will have their own goals to be enabled by grouping and would need to make the decision of how their purpose influences the way the Framework is used. Table 3.2.1 gives some examples of how different purposes might impact on decisions made through the Framework.

Table 3.2.1: Examples of specific purposes for grouping and how this can impact on the way the GRACIOUS Framework is used.

General Purpose	Example	Potential impacts on how the Framework is used
To fill a data gap in a regulatory dossier	A registrant places 5 multiwalled carbon nanotubes (MWCNTs) on the market and needs them to be REACH compliant	All candidate NFs need information to assess a specific hazard endpoint according to REACH. The user will apply the IATA for each grouping hypothesis (for each endpoint), and in the best case the IATA application justifies that some NFs are sufficiently similar, and that one NF is recommended as a *source* material, while all others are target materials. If some NFs are eliminated from the group for a certain pre-defined hypothesis, an attempt to group them using a different hypothesis could be made, or the testing can be escalated to a higher tier method. Measured data needs to be used for the *source* NF and for data gathered via the IATA to support the grouping (and read-across) decision.
To develop precautionary risk management measures	A manufacturer wishes to investigate whether the hazard to their workers from all the TiO_2 NFs they use can be adequately controlled by a single set of risk management measures.	Grouping should target identification of the most hazardous form in the group. If risk management measures control the risk from the most hazardous NF then the risks from all the other NFs should also be adequately controlled.
To steer safe(r)-by-design innovation	The developer of novel paints containing SiO_2 wants to reduce the contribution of this component to the hazard of the complete formulation before committing to optimisation of their coating formulation. Grouping will allow the identification of parameters that they must keep within the NFs of SiO_2 during the development phase.	Consider selection of a range of candidate NFs that are and are not expected to be within the acceptable parameters. For boundaries to be identified it is useful to have examples that fall outside the parameters to properly describe the full extent of a group. These candidates should not be eliminated from the exercise at an early stage as they act as negative controls for the hypothesis.

These impacts are not exhaustive and each user of the Framework will need to identify the impacts specific to their own situation.

3.2.1.2 Worked example: Manufacturer of multiwalled carbon nanotubes (MWCNTs) meeting their REACH obligation

A manufacturer of over 100 tons of MWCNTs per year places 5 different grades on the European Union (EU) market. One grade has sub-chronic inhalation toxicity data (Organisation for Economic Cooperation and Development Test Guideline (OECD TG) 413).

1. **Identify the general purpose of grouping**

The purpose for this particular example is to fill a data gap in a regulatory dossier.

2. **Elaborate the specific purposes of grouping relevant to the user**

At this tonnage level, the REACH regulation requires the provision of long-term repeated dose inhalation toxicity data (Annex IX) for each NF placed on the market in the EU. The user wishes to minimise or eliminate the commissioning of new animal studies while still providing relevant data for this endpoint in their REACH registration. Therefore, they wish to use the NF with existing inhalation toxicity data as the source for read-across to as many other NFs as possible. They also wish to identify whether any NFs might require the commissioning of new *in vivo* studies to be compliant with REACH.

3. **Identify potential impacts of the purpose on the decision-making process through the Framework**

- This endpoint requires a quantitative assessment of similarity to support decision making.
- All justification of the grouping will need to meet the scientific criteria described in the REACH regulation and its supporting documents.
- Data for the endpoint would still need to be provided for any NF that cannot be grouped, before the NF can be placed on the market. This could be accomplished by either read-across from a different group of MWCNTs (e.g. from a different company), or by commissioning new studies. If this is not economically viable, this NF may need to be withdrawn from the EU market.

3.2.1.3 Do you want to know more?

The following resources can provide more information:

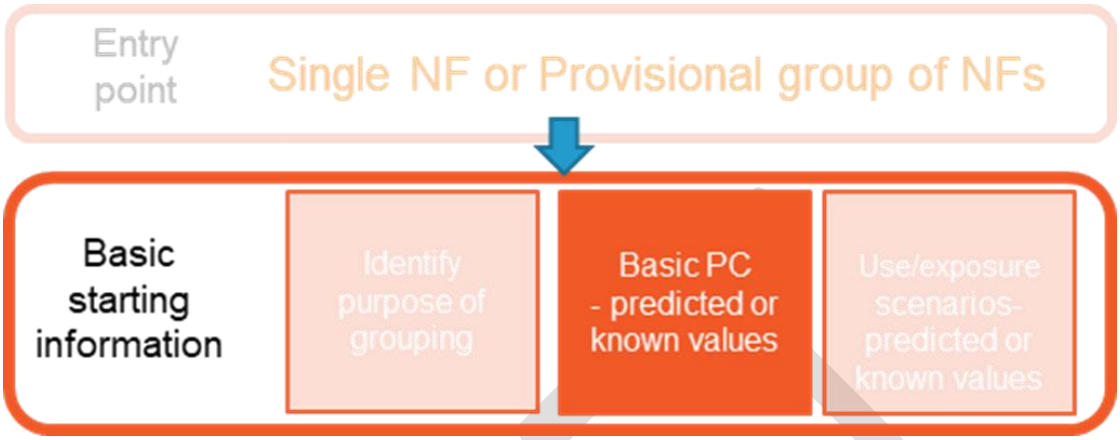
Cooper, R (2017). Idea-to-Launch Gating Systems Better, Faster, and More Agile Leading firms are rethinking and reinventing their idea-to-launch gating systems, adding elements of Agile to traditional Stage-Gate structures to add flexibility and speed while retaining structure. Research-Technology Management, 60, 48-52.

> *Discusses the evolution of the Stage-gate system for efficient product development and launch and introduces how the implementation of aspects of Agile may improve the system even further.*

European Parliament (2006). REGULATION (EC) No 1907/2006 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH).

> *The legal text of REACH states the basic goals of REACH and how these are achieved by requiring defined information requirements on chemical substances placed on the market in the EU on over 1 tonne per annum by a legal entity.*

3.2.2 Physicochemical Characterisation



Physicochemical characterisation provides an understanding of "what they are", which is essential for all aspects of hazard and risk assessment of NFs. This is a principle firmly established in the REACH regulation (Annex VI point 2.3, and nano specific point 2.4), where both chemical and particle characterisation is needed for every NF registered. A set of basic physicochemical characteristics is required in the Basic Information Step of the GRACIOUS Framework, irrespective of the purpose of grouping or the identity of the candidate NFs.

3.2.2.1 Instructions for the measurement of the physicochemical parameters required for the Basic Information of candidate NFs

For SbD approaches and identification of precautionary measures, PC characteristics can be estimated. For progression towards a regulatory application of grouping, estimated values would need to be replaced with either measured values, or if the regulations permit, modelled values (see section 3.2.1).

1. **Identify the required physicochemical parameters**

The revised annexes to REACH now require that a registration dossier should contain the following particle characteristics for every NF covered in the dossier (Annex VI, point 2.4). The GRACIOUS Framework follows these requirements. The list used in the Basic Information step of the Framework is a subset of the particle characteristics required by REACH and is described below. Additional physicochemical characteristics are included in the detailed step of the Framework, but they are IATA specific, so that not all physicochemical characteristics are required for every hypothesis/IATA.

- Composition
    - Identity and concentration of the main constituent and of any impurities.
    - For NFs consisting of an organic substance, Regio isomers, stereoisomers and allotropes should be distinguished and quantified if possible.

- Crystalline phases should be identified and quantified (including amorphous forms). A user can also use the space group number, which together with the chemical it belongs to identifies the 'mineral'. E.g. Anatase => (TiO_2, space group 141).

- Constituent particle size (sometimes incorrectly referred to as primary particle size)
  - In most cases this will be a distribution of sizes. As size distribution may be as important to grouping as median size, a minimum of the D_{10}, D_{50} and D_{90} should be estimated or measured.
  - Depending on the shape of the particles, more than one dimension of the particles may need to be estimated or measured. For example, for a spheroidal particle one dimension can fully describe the size of a NF, whereas for an elongated form, both the width and length will be vital to describe the NF as well as to understand its biological behavior.

Tip for new users

Agglomerate state (size, shape) is dependent on the media in which the NF is suspended, and so it is <u>not</u> included in the basic information. For the Basic Information step, a user should be careful that they are measuring constituent particle size. There is JRC guidance on terminology (2020).

NB. In the detailed step of the Framework, agglomerate state is an important metric in the IATA for some pre-defined hypotheses, so it is worth recording this information if it is measured, but it should always be linked to the composition of the media in which it is measured.

- Particle shape
  - This is a mandatory identifier of a NF under REACH, and specifies the object of investigation. Four broad categories of shape have been suggested by ECHA guidance for REACH registration of NFs (ECHA, 2019) which will be employed here and are defined as follows:
    - **Spheroidal:** particles with an aspect ratio up to 3:1
    - **Elongated:** particles with two similar external dimensions and a significantly larger third dimension (aspect ratio larger than or equal to 3:1). Elongated shape (specifically aspect ratio) is a key parameter used to trigger the HARN IATAs (H-I-1 and H-I-2, see section 3.3.3.2).
    - **Platelets:** particles with one external dimension significantly smaller than the other two external dimensions. The smallest external dimension is the thickness of the particle.
    - **Multimodal shapes:** particles whose shapes belong to different shape categories as the outcome of a manufacturing process and not obtained by mixing particles of different shapes.

Within such generic categories of shape, the ECHA guidance asks for a more precise description of the shape of the particles (so-called shape subcategories e.g. cuboid, wire). For registration, specific information may be applicable such as average **aspect ratio** with an indication of the variation (as a range). However, the GRACIOUS Framework does not request information on subcategories of shape, because these are not relevant in any of the IATA DNs.

The ECHA guidance also uses the term "**assembly structure"**, which include e.g. shell-like structures or hollow structures of constituent particles (ECHA 2019). However, this is not relevant to the GRACIOUS Framework as there is no additional consideration of "assembly structures" in any of the current IATA DNs.

- Chemical nature of the surface
    - The exact nature of any surface modification may depend on the identity of the core NF substance. For example, metal oxides may have pendant hydroxyl groups whose concentration can be engineered and these can be modified by covalent bond formation. Carbon allotropes can display a range of oxygen functionality that may impact on its toxicology (e.g. hydroxyls, carboxylic acids, lactones).
    - Complete characterisation of surface treatment can be difficult, so REACH is satisfied with identification of the reagents (CAS no.) used to *covalently* treat the surface.

Tip for new users

It is often possible to link surface modification with extrinsic parameters of the NF (e.g. surface charge, hydrophobicity). Measurement of these extrinsic parameters are required in the IATA of some pre-defined hypotheses. A user defining their own hypothesis will need to understand the chemical nature of the surface of their NF of interest to be able to identify the appropriate extrinsic parameters to include in their IATA.

- Specific surface area
    - Some studies have indicated that using surface area as a dose metric is more useful than using mass, and ECHA guidance specifically supports assessment of reactivity in surface metric.
    - Specific surface area is required for the Basic Information, and is necessary to evaluate IATA DNs in many of the pre-defined hypotheses.
    - Specific surface area is usually measured using Brunauer–Emmett–Teller (BET) surface area measurement by surface gas adsorption.
    - In some situations, it may be useful to describe internal structures such as porosity.

2. Identify the most appropriate technique for each property

There are often a number of techniques that can be used to measure a given property. The advantages and disadvantages have been well investigated by a number of projects and tools produced to help a user identify the best one for their product. As the principle of similarity (Section 3.3.5) is often used to reach a conclusion on grouping it is very important to use the same technique with all candidate NFs if possible. When choosing the best technique, it is important to consider the following issues:

- Does the technique measure constituent particle size or agglomerate/aggregate size?
- Has sample preparation changed the NF and if so, does this matter? For example, powerful sonication can disrupt agglomerates and maybe even aggregates, thus generating easier access to constituent particles (which is the purpose), while addition of suspension stabilising agents can change the surface chemistry.

Tip for new users

There is extensive guidance available that can support you when choosing the technique most appropriate to your NFs from ECHA, the Joint Research Council (JRC), various EU funded research projects and the International Organisation for Standardisation (ISO). For more details see section 3.2.2.3.

3. Perform the characterisation

Each candidate NF must be characterised according to the list provided above to allow it to be assessed in the Detailed Step of the Framework. The user should remember to make an assessment of any uncertainty in the results in accordance with the guidance on individual methods.

4. Collate the results

Each IATA for a pre-defined hypothesis has its own data matrix (section 3.3.3.1) that allows a user to identify data gaps and to compare results from different candidate NFs. The Basic Information is an intrinsic part of all data matrices. However, at this step of the Framework there is not a formal data matrix so the user should collate the basic physicochemical data, for example, as shown in Table 3.2.2.

Table 3.2.2: Blank table showing the basic physicochemical requirements for the Basic Information in the GRACIOUS Framework. The mandatory requirements are outlined in red.

Property	Method	Unit	NF1	NF2	NF3
Specific surface area					
Constituent particle size distribution					
Composition and impurities					
Surface treatment (CAS #)					
Particle shape					
Crystallinity					
Optional methods					

3.2.2.2 Worked Example: Physicochemical data to meet the Basic Information requirements for five NFs

1. **Identify the required parameters**
2. **Identify the most appropriate technique for each property**
3. **Perform the characterisation**
4. **Collate the results**

As the data matrix collates the information gathered through each step, the worked example is demonstrated by the use of a table generated by the GRACIOUS project (Table 3.2.3).

Table 3.2.3: Summary of the physicochemical data required for each candidate NF in the Basic Information step of the GRACIOUS Framework, using the example of the Representative Test Materials, which can also serve as a proficiency test of the users' laboratory (mandatory parameters outlined in red).

Property	Method	Unit	CNT NM402	CeO2 NM212	BaSO4 NM220	SiO2 NM200	ZnO NM110
Specific surface area	BET	m^2/g	161	27.0	37.0 ± 5.7	190-220	12.0
Constituent particle size distribution	TEM	nm	D10 (width): 5 D50 (width): 10.0 D90 (width): 25 D10 (length): 1200 D50 (length): 1400 D90 (length): 1600	D10: 8.50 D50: 12.42 D90: 20.50	D10: 10.70 D50: 15.50 D90: 30.60	D10: 9,50 D50: 12.50 D90: 15.70	D10: 15.6 D50: 70.0 D90: 105.0
Composition and impurities, option ICPMS	ICP-MS >1000 ppm elements	-	n.d.	Ce	Ba, Sr	Na (1,3%), S (0,8%)	Zn
	ICP-MS 1-1000 ppm elements	-	n.d.	Ni, Pb	Ni, Pb,	Al, Ba, Ca, Ce, Cr, Cu, Fe, K, Mg, Mn, Sr, Ti, V, Zn, Zr	Ni, Pb
Composition and impurities, option XRF	XRF 1-100 %	-	none	Ce (90%)	Ba (60%), S (10%)	SO3, Na2O	n.d.
	XRF 0.1-1 %	-	Al, Si, Fe	none	Al	Cl, Al2O3	n.d.
Composition and impurities, option TGA	TGA	% mass loss of water; of organics	0.3% ; 92.52%	0% ; 0.97%	0.99% ; 3.12%	6.4% ; 4.47%	0;2% ; 89%
Surface treatment	CAS # of surface treatment	-	none	none	none	none	none
Particle shape	TEM	-	Bundle of fibres	spherical	spherical	spherical	spherical
Crystallinity	XRD	-	MWNT with small quantity of impurities	cerianite, cubic	crystalline, orthorhombic	amorphous	Hexagonal, crystalline
Optional: measured surface chemistry: Charge	IEP	pH	3.8	7.5 (0.5)	3.5 (0.5)	3.3	8.5 (0.5)
	Z-potential	mV	n.d.	pH 4: +35.5 (0.1) pH 7: +4.5 (1.3) pH 9: -27.5 (1.9)	pH 4: -8.8 (1.1) pH 7: -35.9 (1.1) pH 9: -38.1 (0.4)	pH 4: -16.0 (1.6) pH 7: -37.4 (0.5) pH 9: -42.0 (1.3)	pH 4: pH 7: +5.3 (0.3) pH 9: -8.1 (4.8)
Optional: measured surface chemistry: hydrophobicity	Water Contact angle	Degree	71.1°	60°	10°	<10°	<10°

Optional: measured surface chemistry:	XPS	Mass%	Pure C (graphite-like)	C 79.9%, O 17.2%, Ce 2.4%	O 52%, Ba 13%, C 17%, S 11%, Cl 3%, P 3%, N 1%	O: 71.43%; Si: 20.30%; C: 5.96 %; Na: 1.83 %	O: 38%; Zn: 35%; C: 30%; Cl: 3%; Na: 3 %
Supporting parameter required by some IATAs: density	He psychometry	g/cm³	2.07	7.12	4.13	2.19	5.67

In the worked example in table 3.2.3, NFs of different substances were assessed, and so different methods were appropriate to measure composition for different NFs. However, it was possible to use a single method, TEM, to analyse both particle size distribution and shape within the same experiment.

3.2.2.3 Do you want to know more?

The following resources can provide more information:

ECHA (2019). Appendix for nanoforms applicable to the Guidance on Registration and Substance Identification. Available at https://echa.europa.eu/documents/10162/13655/how_to_register_nano_en.pdf/f8c046ec-f60b-4349-492b-e915fd9e3ca0

> *This guidance document outlines the particle characterisation required for nanoforms under REACH and suggests methods that can be used to satisfy these requirements.*

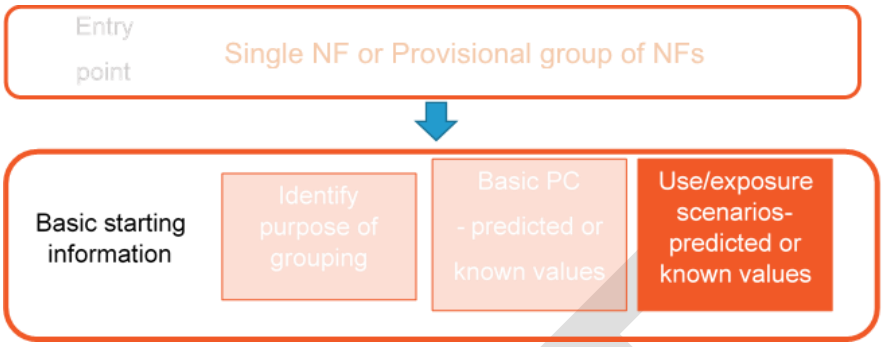
Mech, A., Rauscher, H., Babick, F., Hodoroaba, V., Ghanem, A., Wohlleben, W., Marvin, H., Weigel, S., Brüngel, R., Friedrich, C., Rasmussen, K., Loeschner, K. and Gilliland, D. (2020). The NanoDefine Methods Manual, EUR 29876 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-12336-1, doi:10.2760/58586, JRC117501. Available at https://publications.jrc.ec.europa.eu/repository/handle/JRC117501

> *The NanoDefine Methods Manual has been developed within the NanoDefine project 'Development of an integrated approach based on validated and standardized methods to support the implementation of the EC recommendation for a definition of nanomaterial'. The manual aims to provide guidance through the nanomaterial characterization process, on the use of the characterization methods as well as their application range and their limits to assist the user to choose the most appropriate measurement method(s) to identify any substance according to the EC recommendation for a definition of nanomaterial.*

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

> *Review of new tools to enable regulatory risk assessment of nanomaterials looking at reliability and regulatory relevance as part of the ProSafe project.*

3.2.3 Use and Exposure Scenarios



Having a clear understanding of how NFs are (going to be) used, whether release may occur and how their physicochemical characteristics might change through the life cycle is a vital part of the Basic Information in the GRACIOUS Framework. It will allow the user to identify both the relevant NFs and routes of exposure that require most urgent attention.

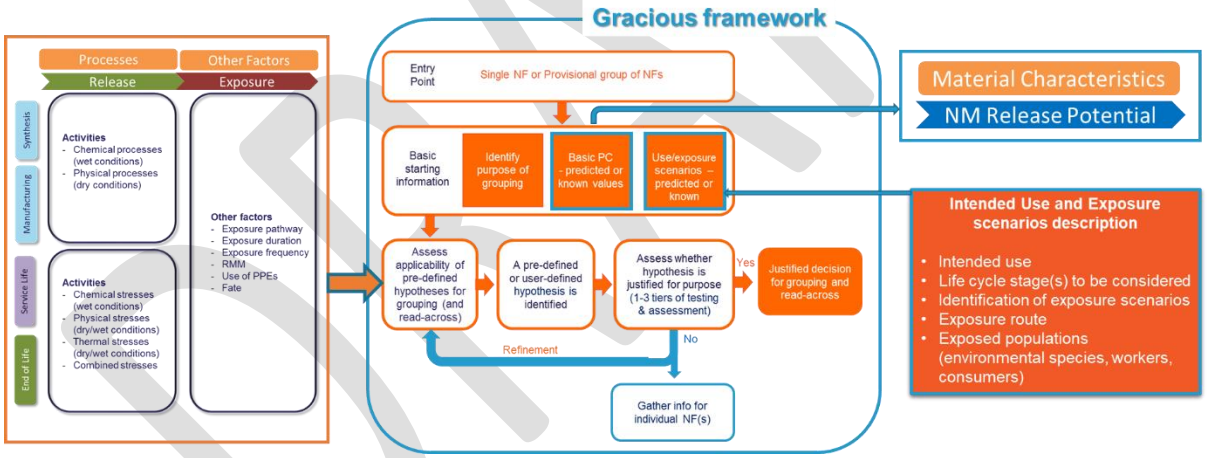


Figure 3.2.1: Representation of the GRACIOUS Framework, highlighting the main input and output related to use, release and exposure

Figure 3.2.1 depicts the GRACIOUS Framework (in central position) with its main components and shows the main inputs (and outputs) related to use, release and exposure to NFs.

The description of the intended uses of the NFs or nano-enabled products and related exposure scenarios in relevant life cycle stages (LCS) will facilitate the identification of the possible release/ exposure paths, the target environmental compartment where the release may occur and the (eventually) exposed populations (workers, consumers, environmental organisms). The description of these variables will guide the user towards the decision tree that might be applicable for each case. It is important to remember

that a risk assessment for a NF might cover impacts on both human health (workers and consumers) and the environment.

The exposure assessment in the Framework can be viewed as a two-stage assessment providing input to different parts of the Framework.

Terminology

Users unfamiliar with REACH may not recognise some of the terms used in this section. Simple definitions for the most important terms are given in table 3.2.4 if more details are needed, please read the guidance documents produced by ECHA: Chapter R-12 Use description and Guidance for downstream users.

Table 3.2.4: Terminology associated with the lifecycle and exposure scenarios of substances

Use	Any processing, formulation, consumption, storage, keeping, treatment, filling into containers, transfer from one container to another, mixing, production of an article or any other utilisation. REACH 2018, Article 3, point 24.
Contributing activities (CA)	Activities contributing to one use. Several activities may take place under one use, leading to several contributing scenarios under one exposure scenario. The contributing activity should be linked to a standardized use descriptor category (e.g., PROC, ERC, PC, AC).
Exposure scenario (ES)	For an identified use (or a group of uses) describes the conditions under which a substance can be used whilst controlling risks. Different contributing exposure scenarios (CES) can be covered under an exposure scenario.,
Contributing exposure scenario (CES)	Specific exposure scenarios associated with a specific contributing activity.
Life cycle stage (LCS)	Stages of the life cycle of a substance. There are four stages to which a use can be assigned: manufacture, formulation or repacking, end-use (including Use at industrial site, Widespread use by professional workers, and Consumer use) and (article) service life.
Process category (PROC)	Describes the tasks, application techniques or process types defined from the occupational perspective, including use and processing of articles by workers.
Environmental release category (ERC)	Describes the activity from the environmental (release) perspective. One ERC is assigned to one contributing activity (environmental perspective) but it can be linked to one or more contributing activities from an occupational perspective (e.g. several PROCs per ERC).

3.2.3.1 Instructions to assess use and exposure

This process can be split into two stages, firstly to identify and describe the uses of the NFs, and secondly to assess the potential exposure during these uses.

Use, Release and Exposure Stage 1: Description of uses, activities and exposure scenarios of NF and NEP

This assessment will define the uses, the activities which contribute to release, and the corresponding exposure scenarios where exposure of humans or the environment is possible. It is recommended to follow the "Use Descriptor" approach described by ECHA for REACH registrations which should initially assess the whole life cycle unless the purpose of grouping specifies a restricted set of scenarios. In combination with the other aspects of the Basic Information Step of the Framework, it will allow the user to identify relevant pre-defined hypotheses or to outline the parameters of a user-defined hypothesis.

1. **Describe the life cycle of NFs and identify the target populations and the environmental compartments affected.**

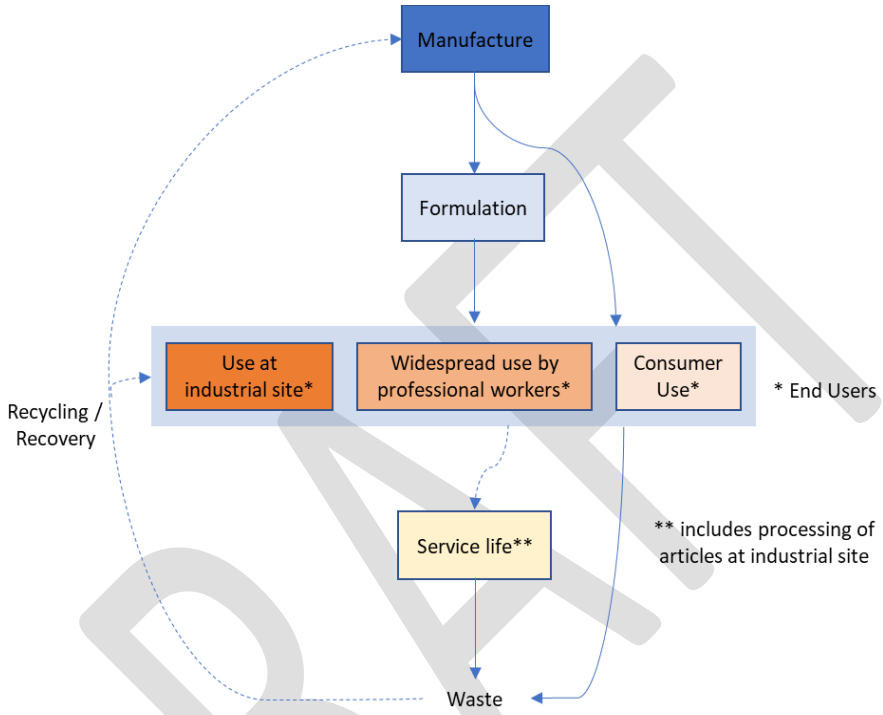


Figure 3.2.2: General diagram showing the life cycle concept (adapted from ECHA R-12 Use description)

Regardless of the purpose of grouping, it is important to identify all the scenarios with potential for release of, and/or exposure to NFs. Following the approach suggested by ECHA (ECHA R-12 Use descriptors), we recommend dividing each use into different activities and to define for each contributing activity (CA) at least one corresponding contributing exposure scenario (CES), considering both human and environmental health (Figure 3.2.2). This will ensure that a common approach is used in generating a structured description of the scenarios and activities that will take place across all life cycle stages, allowing the comparison of release and exposure scenarios and the identification of the uses with highest potential release and/or exposure. Care should be taken when describing scenarios associated with waste life cycle stages. They are currently included in a REACH risk assessment, but the level of detail required is lower than other stages. As regulations adapt to introduce sustainability into their remit, it might become more important in the future (European Commission, 2020).

Also, in this context the use of the ECHA descriptors can provide the user with some intrinsic information contained in the descriptors selected. For example, the selection of the correct use descriptors (e.g., LCS, Process category (PROC), Environmental release category (ERC), etc.) will inform the user on which target

population(s) may receive release/exposure (e.g., workers, consumers, environment species) and the environmental compartment affected (e.g., indoor air, outdoor air, etc.).

Already at this stage, identification of the specific activity and exposure scenario, will allow the user to decide which type of exposure is of most concern:

i. Human vs Environment.

ii. Human exposure: Worker vs Consumer vs General Population.

iii. Environment exposure: Air vs Aquatic vs Soil.

2. Identify uses with similar release profiles

It may be possible to group different exposure scenarios together for the purpose of estimating release and thus simplifying the grouping or allowing a single grouping exercise to be applied across multiple uses. In this context, the use of the ECHA descriptors provides the user with some intrinsic information. For example, the use of NFs in cosmetics and detergents will fall under the same Environmental Release category (e.g., ERC11b Widespread use of articles with high or intended release (indoor)), so it might be possible to use one grouping assessment for both uses. Care must be taken that there truly is similarity between both exposure scenarios, so a user of the Framework should consider the following issues:

- The same NFs are released in each exposure scenario.
- The matrices of nano-enabled products either are similar, or do not impact on the toxicology of the NF.
- Release volumes are similar. Please note REACH does allow comparison of scenarios with some difference in release parameters by the use of "Scaling". If the user is interested in this topic, they should refer to the document "*Guidance for Downstream Users*" published by ECHA (ECHA, 2014).

Use, Release and Exposure Stage 2: Assess likelihood of release/exposure and the physicochemical form of NF during release/ exposure

The next level of description to be provided (or available from GRACIOUS databases) by the user will be related to the Contributing Activities identified for each intended use occurring in the corresponding Contributing Exposure Scenarios. Within these life cycle and Use descriptors, the user needs to understand the likelihood of release/ exposure and the physicochemical form of NF during release/ exposure. For the inexperienced user, the GRACIOUS project has designed decision trees (project deliverables from WP2) that will help with this activity, while experienced users, particularly those familiar with REACH, may be able to perform this task using their own knowledge.

Whilst this information can be used in the Basic Information step of the Framework when identifying a suitable grouping hypothesis, it can also be used to link the grouping conclusion on hazard with the purpose and outcome of grouping, particularly if a risk assessment is the ultimate goal of grouping.

1. **Identify the likelihood of release and the physicochemical form of the NF during release in relevant scenarios**

The GRACIOUS project has made in-depth assessments of some activities that would be regarded as having a high potential for release. From this work, we have developed different decision trees based on activities by taking into account the parameters affecting release from a specific activity. Some of these factors are, for example, related to the NF/nano-enabled product release potential (e.g. the physicochemical form of a NF, or the location of the NF in a nano-enabled product) and to the activity release potential (e.g., type of activity, energy level, etc.). This work allows a better identification of both the activities with highest potential release and the different activities with a similar release potential.

Although the candidate NFs will have been previously identified, once the exposure scenarios and activities of concern have also been identified, it is useful to assess exactly which NF is being released and whether this will be the exact same NF that the target will be exposed to. If the target is exposed to a weathered NF rather than the pristine NF manufactured at the start of the life cycle, it would be important to ensure the release/ exposure relevant NF is indeed in the list of candidate NFs.

2. Identify the likelihood of exposure

The link between release of and exposure to NFs can be interrupted by the use of risk management measures such as operating conditions, technical measures and personal protective equipment. Under REACH, these are only associated with industrial or professional life cycle stages and are not considered for consumer uses or environmental exposure (risk management measures for these targets control release rather than exposure). Therefore, this section is of relevance to exposure in occupational settings. In addition, since release and exposure are sequential events (exposure is not possible if no release occurs), we indicate that when release is likely (for the occupational activities), the user can proceed with an assessment of the likelihood of exposure and the physicochemical form of the NF during the exposure.

The decision tree for assessing likelihood of exposure in occupational settings will depend on system-dependent parameters (e.g., enclosure, local exhaust ventilation etc.) and exposure factors (e.g., exposure pathways, use of personal protective equipment, etc.), including for example the physicochemical form of the NF during the activity and the energy of the activity/process. By answering the Decision Nodes (DNs) of the decision tree, the user will get a conclusion on the likelihood of exposure.

3. Identify the physicochemical form of the NF during exposure

If exposure is likely, the user can proceed with a decision tree to identify the physicochemical form of the NF during the exposure. In fact, also in an occupational setting the physicochemical form of the NF may change from the moment the NF is released to when the target population is exposed. This change is due to the specific conditions (e.g., background particle concentration and size, room size and ventilation, use of enclosure or LEV, etc.) encountered by the NF in the working environment. However, if information on the specific conditions are not available from the user, the physicochemical form of the NF during the release can be assumed to be the same during the exposure.

A worked example is provided (section 3.2.3.2). For this decision tree, the information to be provided by the user concerns the rigidity and needle-like morphology of the NF and the concentration and size of background particles. Based on this information, the decision tree will provide a conclusion regarding the physicochemical form of the NF during the exposure.

4. **Refine exposure assessment using existing tools**

Many other research projects have identified exposure models for nanomaterials that can be applied in different stages of the life cycle. The GRACIOUS Framework encourages the use of these either alongside or instead of the specific decision trees. For example, for the life cycle stages of Manufacturing and Use at industrial site, different occupational models (e.g., NanoSafer http://nanosafer.org/, Stoffenmanager Nano https://nano.stoffenmanager.com/) can be used to obtain a qualitative or semi-quantitative estimation of the worker exposure.

3.2.3.2 Worked Example: Coating of garments with NF Ag

In order to demonstrate how the steps recommended above can be applied, a worked example is outlined below.

Use, Release and Exposure Stage 1: Description of uses, activities and exposure scenarios of NFs and nano-enabled products

1. **Describe the life cycle Use of the NFs and identify the target populations and the environmental compartments affected.**

Firstly, the full lifecycle of a Ag NF, from synthesis to waste is described (Figure 3.2.3).

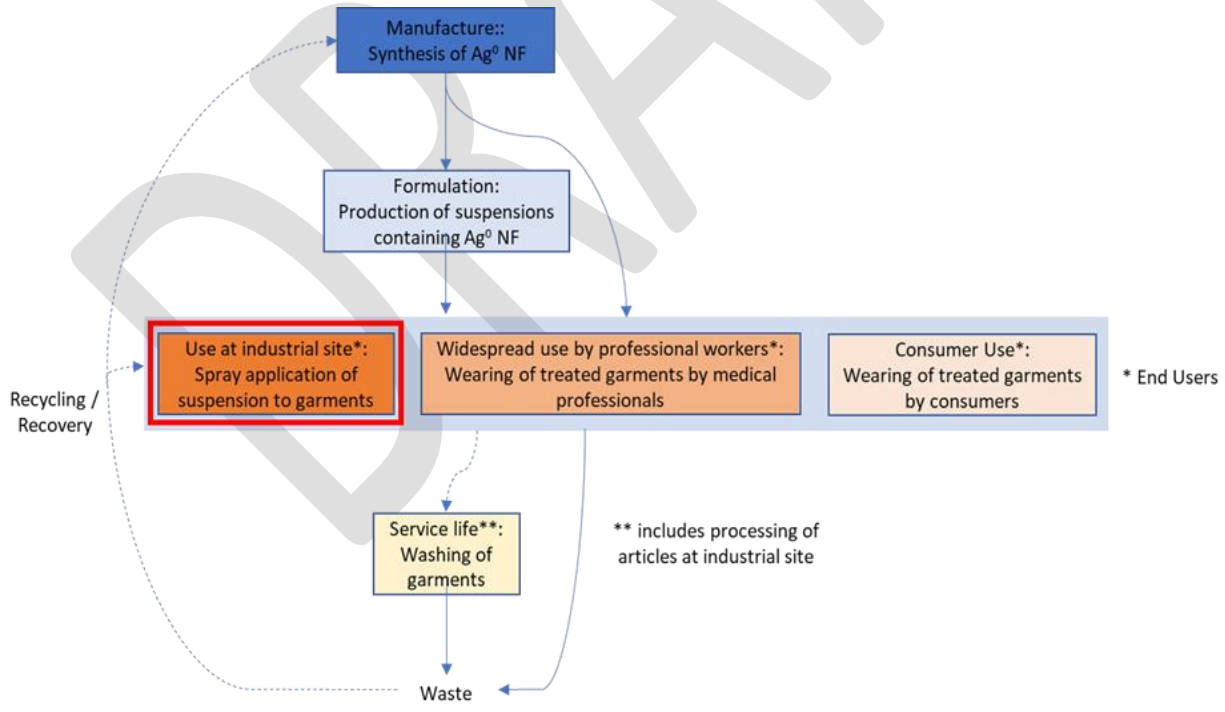


Figure 3.2.3: Diagram showing the life cycle concept (adapted from ECHA R-12 Use description), the red rectangle is highlighting specific LCS of Use at an Industrial site on which the working example is focused.

The user will then identify the uses that need the most attention. This will depend on the purpose of grouping. A REACH registration would need to assess all life cycle stages and contributing activities, whereas other purposes may only need to focus on the uses with the highest potential for exposure. In this Worked Example, the purpose is to understand the risk to the garment manufacturing facility, so the focus is on the "Use at Industrial Site" life cycle stage.

The next step is to fully describe the relevant life cycle stage using Contributing Activities and Contributing Exposure Scenarios. In the worked example, the Use at Industrial Site life cycle stage is detailed in Figure 3.2.4.

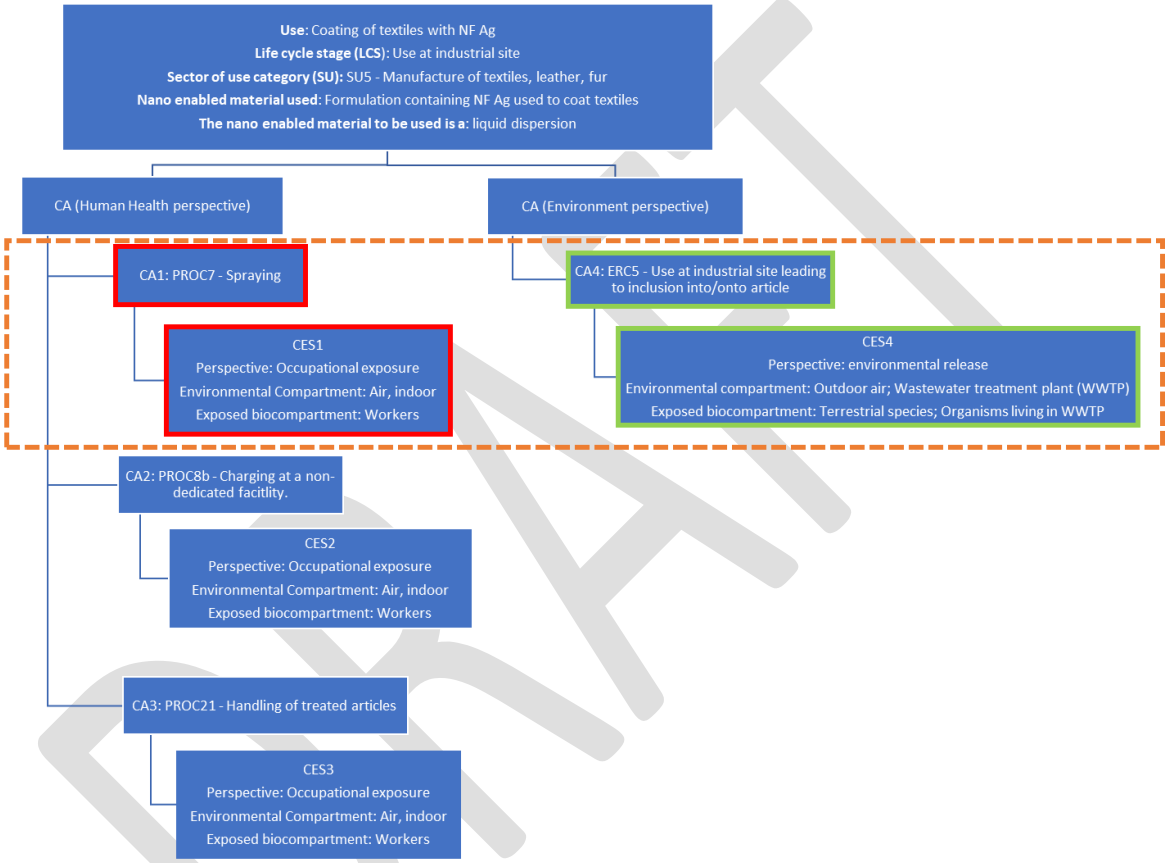


Figure 3.2.4: Schematic representation of different Contributing Activities (CA) and Contributing Exposure Scenarios (CES) from the human health and environmental perspective. The box in the orange dashed line indicates the specific Contributing Activities on which the following explanation will focus on.

Although the Use contains a number of Contributing Activities, the purpose of the grouping requires the user to only assess the one with the highest potential for exposure. For this example **CA2:PROC7 – Spraying** is the one that will be assessed in the rest of the exposure assessment.

2. Identify uses with similar release profiles

As the purpose of the grouping requires the user to focus on the Contributing Activity with highest potential exposure, this step is not needed in this example.

Use, Release and Exposure Stage 2: Assessment of the likelihood of release/exposure and the physicochemical form of the NF during release/ exposure within a specific Contributing Exposure Scenario

The example below focuses on the likelihood of release and exposure and the physicochemical form of the released NF and/or exposure relevant NF in the Contributing Exposure Scenario "Spraying of dispersion containing NF Ag on garment" from the Human health perspective. The example uses decision trees developed in the GRACIOUS project, although the user can use their own approach to mapping release, exposure and physicochemical form.

1. **Identify the likelihood of release and the physicochemical form of NF during release**

In this example, a suspension of the silver NP in an organic/aqueous mixed media is sprayed onto the garments manually by using an atomizer. The key parameters of the process are shown in Table 3.2.5.

Table 3.2.5: Key process parameters for the application of Ag NP to garments used to identify likelihood of release of NF

Parameter	Process characteristic
Level of Energy	High
Presence of organic compounds	Present in suspension/dispersion
Viscosity	Lower than viscosity of water

The relevant decision tree to identify the physicochemical form of the silver NP that may be released is shown in Figure 3.2.5.

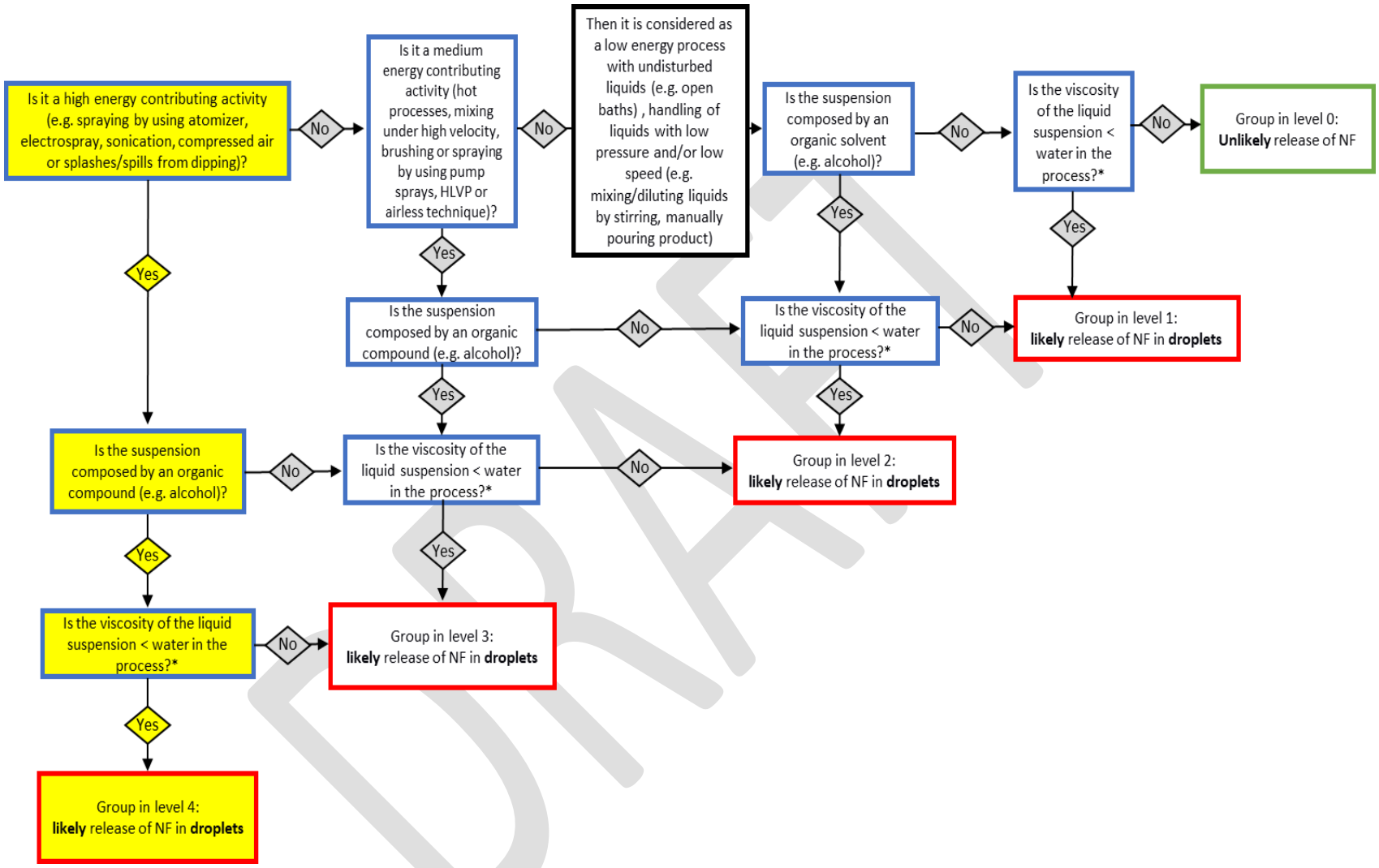


Figure 3.2.5: GRACIOUS decision tree identifying the likelihood of release and physicochemical form of silver NP sprayed onto garments. The route through the decision tree for this process is highlighted in yellow.

Worked example conclusions:

Release of NF is **Likely**

Physicochemical form of NF during release: **droplets**

The substance is released as droplets. There is no evidence that the chemical composition or particle characteristics of released NFs would be different to the manufactured NF. There is no information on the impact that being suspended in a liquid media has on the agglomeration behavior of the NF.

2. Likelihood of exposure

As the contributing activity is an industrial life cycle stage, risk management measures can be used to prevent the exposure of workers to released particles. The process parameters for this worked example are shown in Table 3.2.6.

Table 3.2.6: Key process parameters used to estimate likelihood of exposure of workers to Ag NP during application to garments.

Parameter	Process Characteristic
Segregation of emission source and workers	No segregation
How is the process enclosed?	Full containment
Containment procedure	Access points closed
Further technical measures	Local Exhaust Ventilation (LEV) used
Personal Protective Equipment	Full face mask

In this example, a GRACIOUS decision tree is used to understand the likelihood of exposure (Figure 3.2.6) but the user can use their own knowledge of their process to assess it.

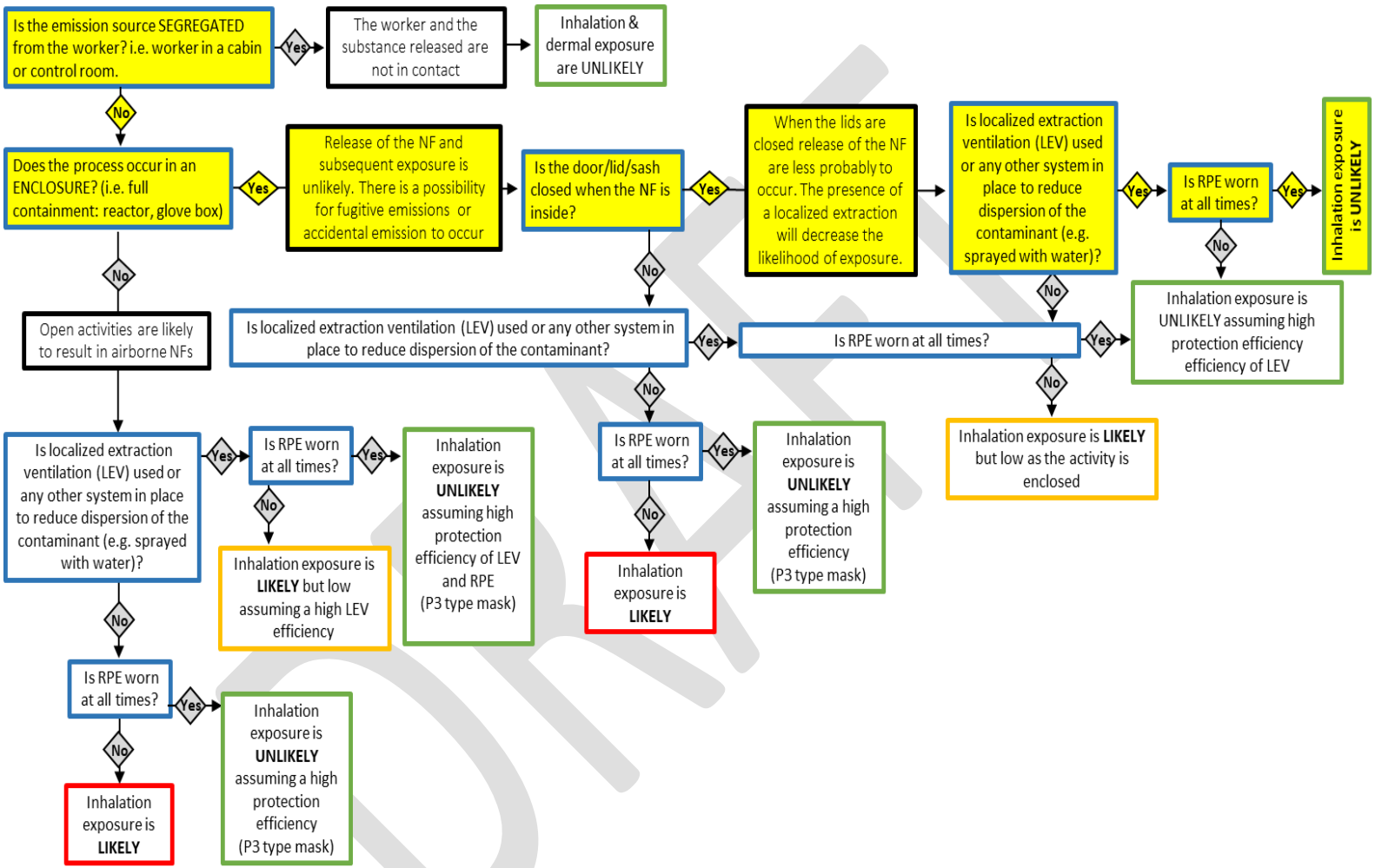


Figure 3.3.6: Decision tree to identify likelihood of exposure toAg NP during spray application to garments. The route through the decision tree for this process is highlighted in yellow.

Worked example conclusion:

Inhalation Exposure is **Unlikely**

The user needs to decide how to use this information. If grouping is being considered for regulatory purposes, hazard assessment is obligatory so the process would need to continue with an emphasis on inhalation exposure. However, if the purpose of grouping is to identify adequate risk management measures, it may be possible to conclude that if the existing measures are in place, no further assessment is needed because exposure to workers is unlikely.

3. Physicochemical form of NFs during exposure

For a risk assessment to be truly relevant it should include the NFs to which the worker is exposed. The key physicochemical parameters used to make this judgement for this worked example are shown in Table 3.2.7.

Table 3.2.7: Key process parameters used to estimate physicochemical form of NFs during the application of Ag NFs to garments.

Parameter	Process characteristic
Morphology	Spherical
Background particle concentration	3E+06
Background particle diameter (mean)	200 µm

The GRACIOUS decision trees are used to identify the physicochemical characteristics of the NF that workers will be exposed to (Figure 3.2.7), but the user can also use their own experience and information.

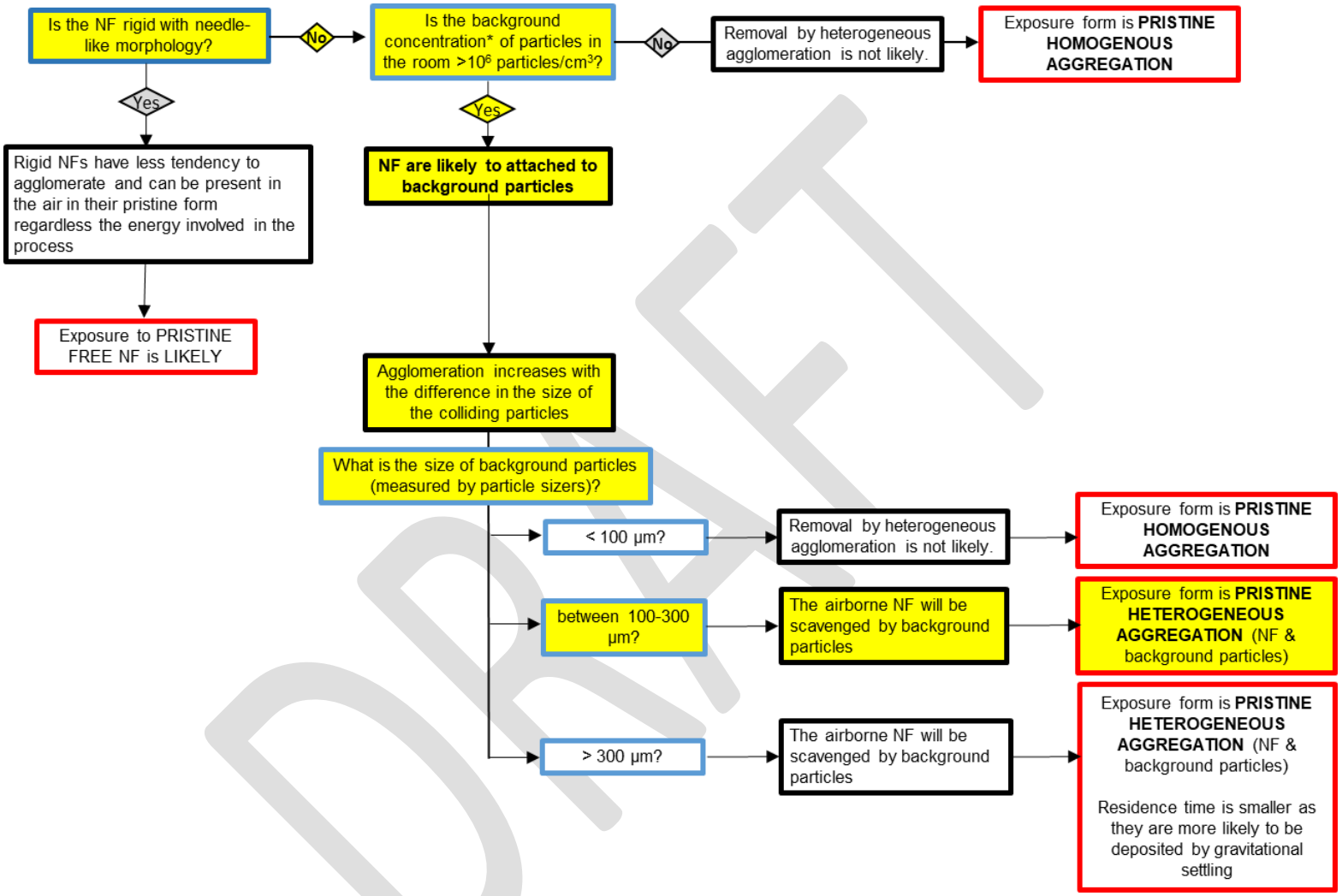


Figure 3.2.7: GRACIOUS decision tree to identify the physicochemical characteristics of Ag NFs that workers may be exposed during the spraying of garments.

Worked example conclusion:

Physicochemical form of NF during exposure: **Heterogeneous aggregation/agglomeration of pristine NF and background particles**

If the purpose of the grouping is for REACH registration, endpoint data needs to be provided on pristine forms, but if a risk assessment is required the registrant would need to assess whether the aggregates /agglomerates of pristine and background particles should be a candidate NF or whether it can be argued that pristine NFs will be more toxic than the aggregates/agglomerates with background particles and hence data on the pristine forms can be used in the risk assessment.

3.2.3.3 Do you want to know more?

The following resources can provide more information:

European Commission (2020). Chemicals Strategy for Sustainability Towards a Toxic-Free Environment. Available at https://ec.europa.eu/environment/pdf/chemicals/2020/10/Strategy.pdf

> *Outlines the European Commission's goals to achieve a toxic free environment. It is likely to act as the guideline for new regulations or amendments of existing regulations in the EU over the next 5 - 10 years.*

ECHA (2014). Guidance for downstream users. Available at https://echa.europa.eu/documents/10162/2324906/du_en.pdf/9ac65ab5-e86c-405f-a44a-190ff4c36489

> *This guidance document describes the obligations that downstream users (not manufacturers or importers) of substances have under REACH. It explains how a downstream user can prove that the risk management measures they put in place are equivalent or better than those included in the registration dossier of a substance and hence demonstrate compliance with REACH.*

ECHA (2015). Guidance on Information Requirements and Chemical Safety Assessment, Chapter R.12: Use description. Available at https://echa.europa.eu/documents/10162/13632/information_requirements_r12_en.pdf/ea8fa5a6-6ba1-47f4-9e47-c7216e180197

> *This guidance document explains how uses and activities can be described using the Use Descriptor system, allowing all registrants to use the same approach when describing the use of substances across all industries and sectors.*

Project deliverables from Work Package 2

D2.1 set the basis for the input to be provided to the framework, identifying criteria and guiding principles for grouping and read across of NMs/NFs relevant to release and exposure.

D2.2 aimed at providing relevant NF, NEPs and case studies for hypothesis testing in WP2-5.

D2.3 provides an insight on the descriptors, criteria, and guiding principles for grouping and read-across focusing on release and exposure scenarios.

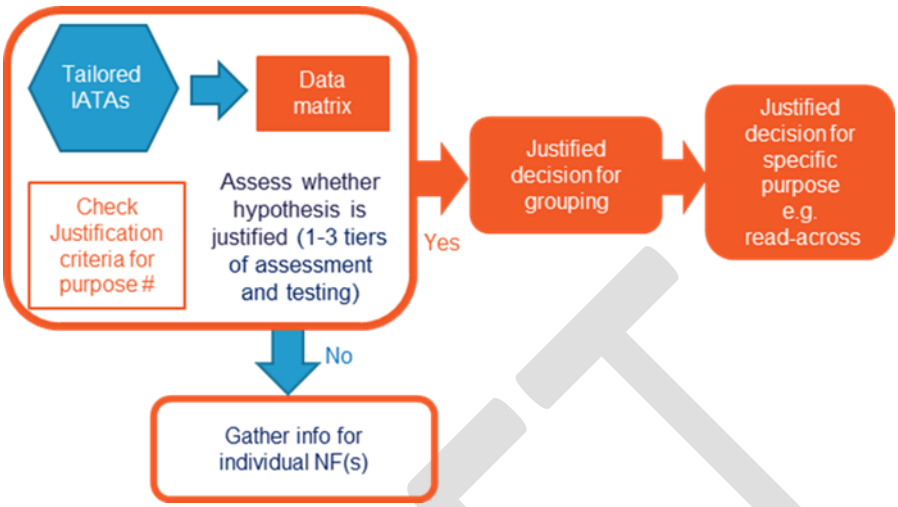
D2.4 includes the generation of two templates (occupational release & exposure (ORE) template and the environmental & consumer release (ECR) template) aimed to organize and collect data on release and exposure to NMs/NFs. The templates will be the basis to develop the Gracious exposure and release library, that will be one of the main contributions of WP2 to the GRACIOUS framework.

D2.5 provided hypothesis, decision trees and corresponding tiered testing strategies. A manuscript on the decision trees is under preparation in WP2.

D2.6 contains experiments, tests, methods and/or tools can be useful to demonstrate the different hypothesis and to support/help the user of the GRACIOUS framework.

D2.7 contains the description of each release/exposure-related component (e.g., use and exposure scenarios, hypothesis, decision trees and tiered strategy) developed for inclusion in the Framework.

3.3 Detailed Step: Testing the Grouping Hypothesis



Once the Basic Information for all candidate NFs is collected, it can be used to define the grouping hypothesis. A number of pre-defined hypotheses have been generated and tested. As much of the background work for these hypotheses has already been done (e.g. building IATAs, identification of suitable testing protocols, definition of decision nodes (DNs)), the user should investigate whether these pre-defined hypotheses apply to their NFs, or whether they can modify them to meet their needs.

3.3.1 Selecting a shortlist of pre-defined hypotheses

A significant amount of work on the fate, behavior and hazards of nanomaterials has been published. The GRACIOUS project has used and expanded this work to develop 44 pre-defined hypotheses along with the development of IATAs to test these hypotheses. The pre-defined hypotheses span all primary routes of exposure to humans and across most of the major environmental compartments. The Basic Information step of the Framework is unlikely to be sufficient to select a single pre-defined hypothesis, but it can be used to identify a shortlist of hypotheses that can be refined by performing some studies in the Detailed step of the Framework.

3.3.1.1 Instructions to select a shortlist of pre-defined hypotheses

In order to identify a shortlist of relevant pre-defined hypotheses, the following process is recommended.

1. **Collate all basic physicochemical data**

The GRACIOUS Framework requires that the basic physicochemical characterisation data is available for all NFs under consideration for grouping. They largely follow the data required by Annex VI of REACH as described in section 3.2.2 of this document.

2. **Identify the route and type of exposure of concern**

The Basic Information on the life cycle of the NFs will allow the user of the Framework to identify the type and route of exposure that will need to be investigated (Guidance is provided in section 3.2.3). The user

should also consider the impact of the purpose of grouping on this choice. For example, if the Framework is being used for regulatory purposes, the regulation itself will specify which routes need to be investigated. For example, a registration of NFs of a substance in the 1–10 tonne tonnage band in REACH requires that acute toxicity via inhalation is tested according to section 8.5.1 of Annex VII, unless exposure via inhalation is unlikely across the life cycle of the NFs.

3. Identify a shortlist of relevant pre-defined hypotheses

The physicochemical characteristics and the route of exposure are used to identify the hypothesis shortlist. At this step, a single hypothesis should not be identified to the exclusion of all others as different NFs display different behaviors in the same route of exposure depending on their physical and chemical characteristics. The pre-defined hypotheses have been designed to cover many of these variations. In some cases, these characteristics are not within the Basic Information requirements, instead their acquisition is guided by the IATA triggered by a pre-defined hypothesis. The flow chart in Figure 3.3.1 shows how the choice of a relevant shortlist of potential pre-defined hypotheses can be made.

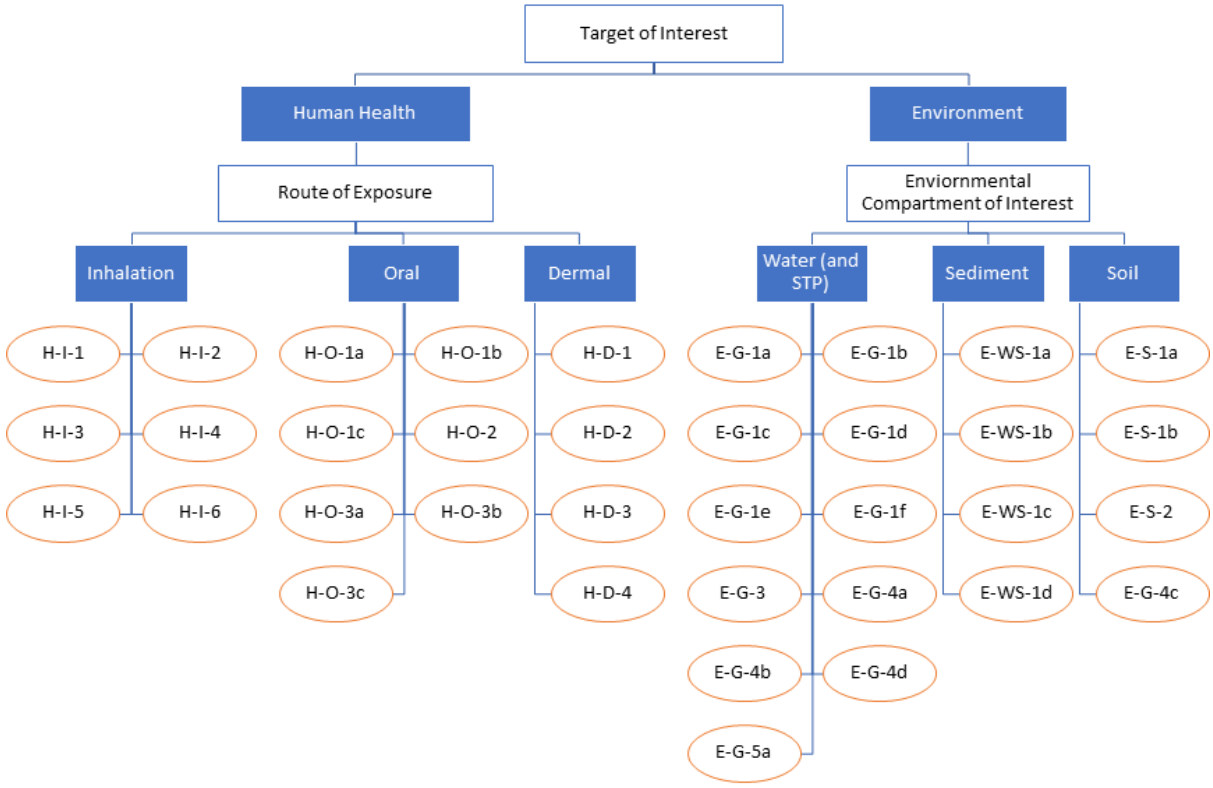


Figure 3.3.1: Flowchart to identify a shortlist of pre-defined hypotheses based on the Basic Information step

4. Consider the consequences of the grouping

Although the result of the grouping exercise cannot be predicted, it may be useful to consider the consequences of the outcome of the grouping exercise. These will align with the purpose of grouping and should include the result of the grouping being fully or partially successful or unsuccessful. For example,

when grouping for the purpose of SbD, a successful or partially successful grouping means all members of a group demonstrate too much toxicity to progress with product development, but any NF that could not be included in a group can be considered for further development. An unsuccessful grouping may require the use of an alternative pre-defined hypothesis or the development of a user-defined hypothesis.

Tips for new users

The advantage of using a pre-defined hypothesis is that the testing and how to interpret the results is already well established by the GRACIOUS project. Data rich NFs that fall within the hypothesis, possibly as part of a pre-existing group, will be available avoiding the need to search databases. The best approaches to assessing similarity will be established and methods to assess the quality of data should be in place (possibly the assessment for any existing data will already have been done).

3.3.1.2 Worked Example: Inhalation toxicity of Carbon Nanotubes

A case study investigating the grouping of five MWCNTs was performed to test one of the pre-defined grouping hypotheses and its associated IATA. The purpose was to provide the data needed to address REACH requirements for inhalation toxicity. This example demonstrates how the most appropriate hypothesis can be selected.

1. Collate all basic physicochemical data

A summary of the Basic Information on the 5 candidate MWCNTs is shown in Table 3.3.1.

Table 3.3.1: A table of the "what they are" Basic Information.

MWCNT	Carbon (%)	Length Mean± SD (µm)	Diameter Mean± SD (nm) (range)	Shape Aspect Ratio (D3:D1)	BET (m^2/g)	Crystallinity
MWCNT-A	86.2	0.85±0.10	11±3 (6-17)	Elongated 77.27	254	Graphenic
MWCNT-B	99.7	4.0±0.37	67±24 (24-138)	Elongated 59.7	18	Graphenic
MWCNT-C	96.1	1.4±0.19	11±3 (7-20)	Elongated 127.27	226	Graphenic
MWCNT-D	99.1	0.4±0.03	12 ±7 (5-37)	Elongated 33.33	135	Graphenic
MWCNT-E	99.6	5.7±0.49	74 (29-173)	Elongated 77.02	26	Graphenic

2. Identify the route and type of concern

The purpose of the grouping is to provide the data needed that would otherwise be generated by repeated dose animal toxicity studies in order to address the data requirements under REACH. The life cycle input demonstrates that aerosolization is possible during production. Therefore, human (worker) hazard is a principal concern and the REACH regulation requires that inhalation is the primary route of exposure to consider.

3. Identify a shortlist of relevant pre-defined hypotheses

Using the information above a shortlist of the relevant pre-defined hypotheses can be identified.

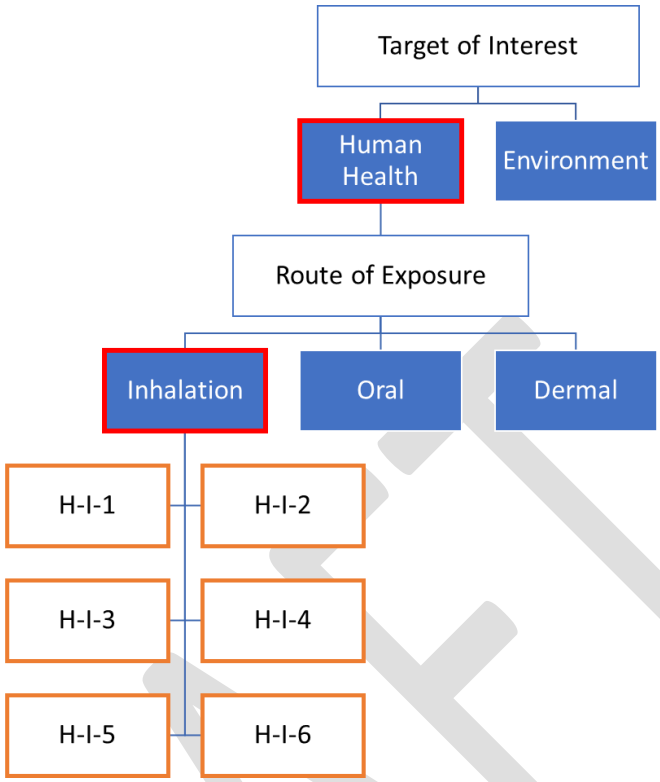


Figure 3.3.2: Demonstration of how the full list of pre-defined hypotheses can be refined to a shortlist for the grouping of different MWCNTs to satisfy REACH requirements for inhalation toxicity to workers.

The group of potentially relevant hypotheses are:

H-I-1	*"Respirable, bio persistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of NFs can occur resulting in lung toxicity."*
H-I-2	*"Respirable, bio persistent, rigid HARN longer than 5 µm: Following inhalation exposure and translocation of NFs to the pleura, mesothelioma development can occur."*
H-I-3	*"Respirable NFs with an instantaneous dissolution: Following inhalation exposure, the toxicity is driven by and is therefore similar to those of the constituent ions or molecules."*
H-I-4	*"Respirable NFs with a very slow dissolution rate: Following chronic inhalation exposure, accumulation of NFs in the lungs can lead to long-term toxicity."*
H-I-5	*"Respirable NFs showing partial dissolution: Following inhalation exposure both NFs and constituent ions or molecules may contribute to toxicity and there is some concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release."*
H-I-6	*"Respirable NFs showing quick dissolution: Following inhalation both NFs and constituent ions or molecules may contribute to toxicity, but there is no concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release."*

4. **Consider the consequences of the grouping**

As the consequences of the grouping and the potential impact of different outcomes will depend on the purpose of grouping, they will vary from case to case. In this scenario, the purpose is to fill data gaps in a REACH registration dossier. This leads to the following potential consequences for the different potential outcomes:

a. All NFs fit into one group
   i. One set of data will be sufficient for the repeated dose toxicity for all NFs.
   ii. All NFs can be considered to have a similar inhalation hazard.
   iii. If this is a significant hazard, strict risk management measures of all NFs may be required.

b. The candidate NFs are divided into more than one group
   i. More than one OECD TG 412/413 studies will be required to cover each group of NFs.
   ii. Will the risk assessment of two groups be cost effective?
   iii. Will some NFs be removed from the EU market due to the cost of risk assessment or a lack of information?
   iv. Will different risk management measures be needed for each group of NFs?

c. Some NFs do not belong in a group
   i. Will *in vivo* endpoint testing be required on each of these non-grouped NFs?
   ii. Will a new hypothesis need to be written and tested?
   iii. Will the non-grouped NFs be removed from the EU market due to the cost of risk assessment or a lack of information?

3.3.2 Refining the shortlist to identify the most relevant pre-defined hypothesis

This section outlines how a user can refine their pre-defined hypothesis shortlist further to identify hypotheses that are applicable to their unique situation. The user must be aware that the Framework gives a structure to using best practices for grouping. However, as introduced in section 3.2.1, the Framework can be used for a wide variety of grouping purposes, so the user must always consider their individual purpose and desired outcomes when using IATAs to identify and test a hypothesis. This section does not give guidance on how to reach a final conclusion on grouping, as this is covered in Section 3.3.5.

3.3.2.1 Instructions for refining the shortlist to a single pre-defined hypothesis

1. **Identify the most probable hypothesis**

To avoid the need to work on several IATAs simultaneously, the user can use their knowledge of the candidate NFs to predict the most probable pre-defined hypothesis. It is important at this step to not eliminate any hypotheses without evidence, as it is possible that the results of the studies in the IATA associated with the probable (or convenient) hypothesis result in some or all of the NFs being rejected from the potential group. If the probable (or convenient) hypothesis is rejected, then one of the other hypotheses could be relevant (this is further explained in the Worked Example). Fortunately, there are often common DNs in the different IATAs, so it may be possible to apply data collected in the initial IATA to any subsequent IATAs that might be required.

2. **Start the IATA of the most probable hypothesis and use results to eliminate other hypotheses in the short-list**

For many of the lists of hypotheses shown in Figure 10, the different hypotheses are defined by the value of a physicochemical (e.g. rate of dissolution) or biological (e.g. inflammation) parameter. Upon measurement of a parameter for all the candidate NFs it will become clear if all, some or none of the NFs meet the criteria of a short-listed hypothesis. It may be that an individual test or DN in the IATA of the probable hypothesis does not help with the refinement of the shortlist. The results of the test or DN will still be crucial to the IATA and the final grouping conclusion, so it is important not to omit them in order to identify a single hypothesis quickly.

Tip for new users

At this stage, it is not essential to test all NFs to the exacting standards of the studies required when using a specific IATA to test a pre-defined hypothesis. It may be possible to use expert judgement or data on a limited number of candidate NFs (e.g. any data rich NFs that may be used as the source for a read-across).

At this stage the aim is to refine the shortlist by removing hypotheses that do not apply to any of the candidate NFs. If a hypothesis appears to address some candidates and not others, it is useful to keep it in the shortlist. Once a pre-defined hypothesis is selected and the IATA started, some candidate NFs can be removed from the proposed group if the experimental results support this.

It is not essential to address all DNs in the IATA, only enough to identify the best pre-defined hypothesis to investigate initially.

3. **Consider options where only some candidate NFs meet the hypothesis criteria**

Where all of the candidate NFs meet the criteria of a short-listed hypothesis, it will be easy to decide to include the hypothesis on the short-list. Similarly, if none of the candidate NFs meet the criteria of the short-listed hypothesis, it will be easy to eliminate such a hypothesis. However, it may be that some NFs in the potential group meet the criteria for a specific hypothesis whereas others do not. In this situation, the user must refer to the purposes and expected outcomes to decide on the next steps to take. Some of the options the user could consider are detailed in Table 3.3.2.

Table 3.3.2: Examples of how the purpose of grouping can be used to decide on the next step when only some candidate NFs meet the criteria of a pre-defined hypothesis

Purpose	Next step	Reason
Scientific research: to identify limiting parameters linked to an adverse apical endpoint result.	Continue with IATA for most probable hypothesis with all candidate NFs.	Negative results are as important as positive results to define the parameter boundaries.
Safe-by-design: Identification of high-risk NFs from product development.	Remove any NFs from candidates that do not meet the criteria for the probable hypothesis.	The goal is to identify the group of NFs that have a high risk to safety due to a particular biological behavior. As soon as a NF can be shown not to display this behavior, further testing with it is unwarranted.
Satisfy a regulatory endpoint.	Continue with the most probable hypothesis for candidate NFs that satisfy criteria for the hypothesis. Consider whether a different short-listed hypothesis is still relevant for candidate NFs that do not meet the criteria.	The intention is to satisfy an endpoint for all candidate NFs. If it is clear that the probable hypothesis is relevant to only a limited number, the endpoint still needs to be satisfied for the other NFs.

Tips for new users

Consider using software tools which contain implemented (parts) of the GRACIOUS Framework based upon the Blueprint to get support in the decisions making process. If fully implemented in the software, the software tool(s) should allow a user to insert their basic information and purpose and it can identify the most appropriate pre-defined hypothesis using this information. It may mean that some aspects of the data matrix will be completed even before starting the IATA, if the software contains this functionality.

3.3.2.2 Worked Example: Inhalation toxicity of Carbon Nanotubes

In the previous section, a worked example using MWCNTs was introduced. The purpose for grouping and the use information were used to identify a shortlist of potential pre-defined hypotheses.

Purpose: Regulatory to fulfil repeated dose toxicity requirements of REACH

Exposure context: Occupational

Life cycle input: Exposure from aerosolization of powder during production of key concern. Inhalation is the key route of exposure

Shortlist of hypotheses:

H-I-1	*"Respirable, bio persistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of NFs can occur resulting in lung toxicity."*
H-I-2	*"Respirable, bio persistent, rigid HARN longer than 5 µm: Following inhalation exposure and translocation of NFs to the pleura, mesothelioma development can occur."*
H-I-3	*"Respirable NFs with an instantaneous dissolution: Following inhalation exposure, the toxicity is driven by and is therefore similar to those of the constituent ions or molecules."*
H-I-4	*"Respirable NFs with a very slow dissolution rate: Following chronic inhalation exposure, accumulation of NFs in the lungs can lead to long-term toxicity."*
H-I-5	*"Respirable NFs showing partial dissolution: Following inhalation exposure both NFs and constituent ions or molecules may contribute to toxicity and there is some concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release."*

H-I-6	*"Respirable NFs showing quick dissolution: Following inhalation both NFs and constituent ions or molecules may contribute to toxicity, but there is no concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release.*

1. Identify the most probable hypothesis

User knowledge and the Basic Information indicates that MWCNTs have a high aspect ratio. HARNs are known to have the potential to trigger significant adverse effects when inhaled. Therefore **H-I-1** was selected as the most appropriate hypothesis to start to investigate grouping for the REACH repeated dose endpoints. It must be noted again that this selection does NOT mean the other hypotheses in the shortlist have been eliminated, as the defined information does not warrant it.

2. Start IATA of most probable hypothesis and use results to eliminate other hypotheses in the short-list

A description of IATAs in general is provided in section 3.3.3, and the full IATA for **H-I-1** is described in Section 3.3.3.2, Figure 3.3.6. In this example, we will only move through the decision nodes (DN) that can allow elimination of other shortlisted hypotheses, and therefore is limited to DN1- DN4. Figure 3.3.3 lists the basic questions asked in each DN of H-I-1 and the criteria associated with the questions.

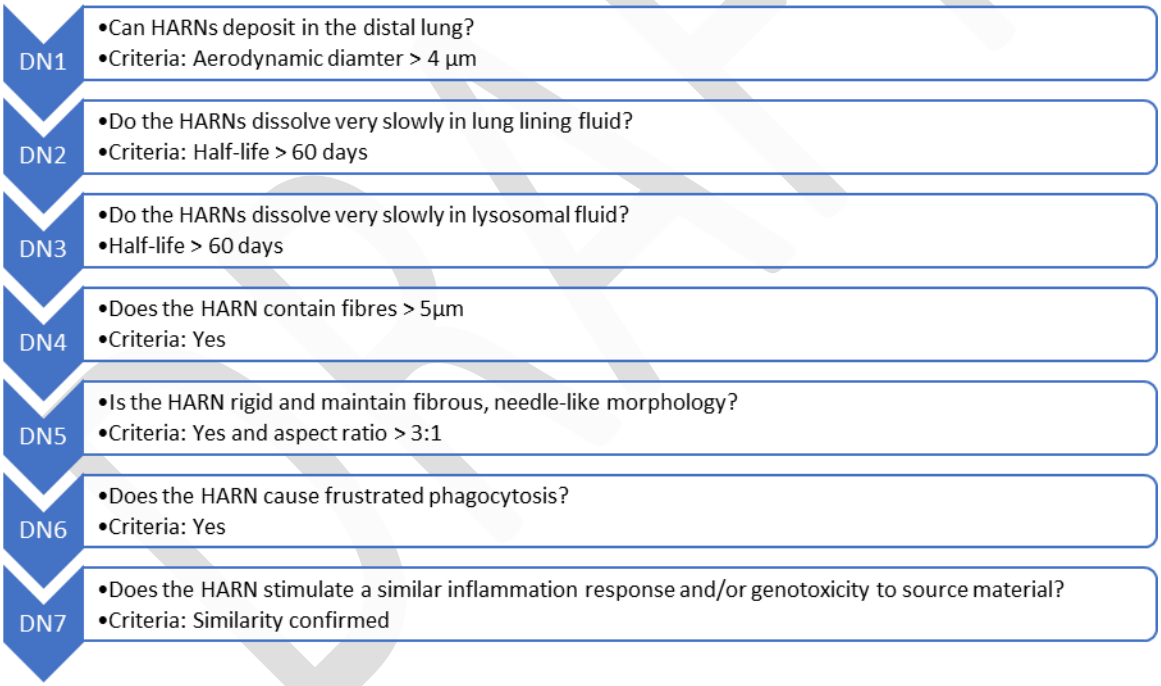


Figure 3.3.3 A simplified description of Decision Nodes (DN) in IATA H-I-1.

DN1: Can NF deposit in the distal lung?

Criterion:	Aerodynamic diameter > 4µm
Can this criterion eliminate other shortlisted hypotheses?	No, this is a criterion for all IATAs of short-list hypotheses
Do all candidate NFs meet the criterion?	Basic information indicates that all candidates meet this criterion.

DN2/3: Does the NF dissolve in lung lining fluid/lysosomal fluid very slowly?

Criterion:	Half-life > 60 days
Can this criterion eliminate other shortlisted hypotheses?	Yes, **H-I-3**, **H-I-5** and **H-I-6** can be eliminated as they require more rapid dissolution. **H-I-2** and **H-I-4** remain in the short-list
Do all candidate NFs meet the criteria?	Yes, expert knowledge indicates that any substances with a graphenic structure will be insoluble in these fluids.

DN4: Are HARNs length > 5um?

Criteria:	10% of HARN fibres > 5um in length.
Can this criterion eliminate other shortlisted hypotheses?	Yes, **H-I-4** does not specify fibre morphology as a criterion for grouping. **H-I-2** would remain in the short-list.
Do all candidate NFs meet the criteria?	Requires detailed analysis of samples at both constituent particle level and agglomerate level which may not be available at Basic Information step. Conservative approach is to assume 'yes' until data is generated within the IATA process.

At this stage **H-I-1** and **H-I-2** are in the shortlist. **H-I-1** is targeted towards lung toxicity whereas **H-I-2** specifically investigates the potential for mesothelioma. As the purpose of grouping in this case is to provide data to satisfy the inhalation toxicity endpoint of REACH, **H-I-1** is the most appropriate hypothesis to investigate initially.

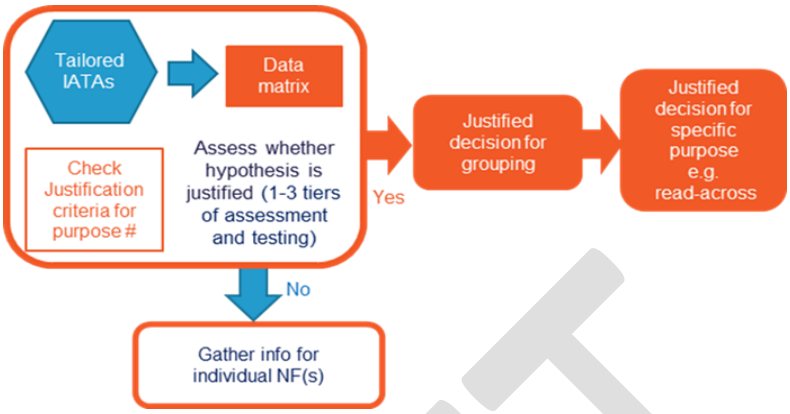
1. Consider options where only some candidate NFs meet the hypothesis criteria

The result at DN4 means that the user needs to consider their next step and this should be guided by the purpose of the grouping. In this scenario, the purpose is to provide the repeated dose toxicity data relevant to inhalation, in line with REACH information requirements. To achieve this the user would use or generate one set of data using OECD TG 412 or OECD TG 413 and apply the results for read-across to the target group members. Potentially a majority of all candidate NFs may not meet the criteria to continue with IATA **H-I-1** (i.e. for these NFs hypothesis **H-I-1** is rejected) and this scenario's purpose will not be achieved. Therefore, the user must decide how to proceed. Two possible options are given below, but these are not exhaustive:

a. The candidate NFs that do not meet the criteria for **H-I-1** may meet the criteria for **H-I-4**. Consider running two parallel grouping exercises using each pre-defined hypothesis.
b. The criteria for rigidity may need investigation for this substance, continue with **H-I-1** studies with all candidate NFs to investigate whether further studies indicate that similarity of rigidity is a better parameter to identify a HARN.

H-I-1 is chosen as the pre-defined hypothesis to test, with the caveat that if strong evidence of the potential for mesothelioma is found for some or all candidate NFs, the work can be extended to **H-I-2**. If some NFs are eliminated from the HARN group (i.e. hypothesis H-I-1 is rejected), they can be examined using hypothesis **H-I-4**.

3.3.3 Using an IATA for a pre-defined hypothesis



Once a pre-defined hypothesis is identified, the GRACIOUS project has designed an IATA as a structured way of identifying the most relevant information to gather to reach a grouping conclusion with strong scientific justification.

The IATA is made up of decision nodes (DNs) that will answer questions vital to justifying a grouping conclusion. These DNs encourage the use of existing data, but where data gaps exist, they guide the acquisition of new data either via experimental or *in silico* studies (Figure 3.3.4).

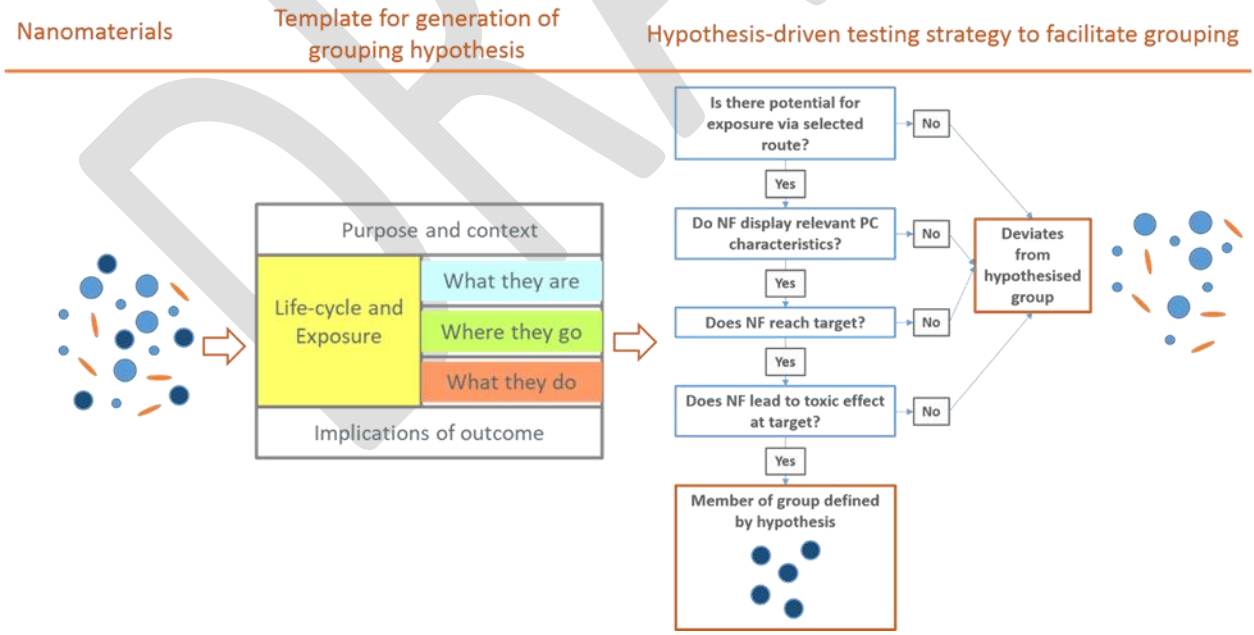


Figure 3.3.4: IATAs are used to define the information that is required to test a grouping hypothesis in the GRACIOUS Framework. The IATA is made up of decision nodes (DN) (blue boxes).

The study designs or protocols recommended at each DN are organised into three tiers of increasing complexity, although not every DN will have three tiers. For human hazards, tier 1 generally consists of *in silico*, physicochemical or simple *in vitro* studies, tier 2 uses more complex *in vitro* studies and tier 3 requires *in vivo* studies (Figure 3.3.5). For environmental hazards, tier 1 will include acute hazard assessments using standard species (e.g. Daphnia magna), tier 2 will include longer-term hazard studies, while tier 3 might include mesocosm studies. It may not be necessary to proceed through all the tiers to reach a decision on how to progress through the DN. This depends on how conclusive the results are and the purpose of the grouping exercise.

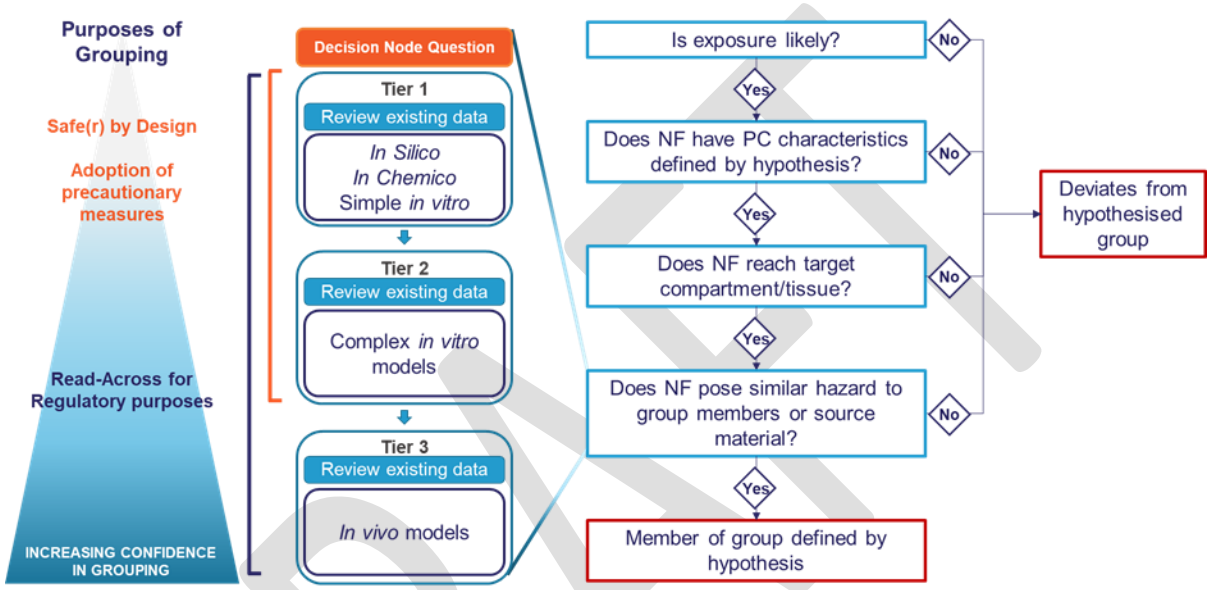


Figure 3.3.5: Generic structure of a Decision Node showing how the tiered structure can be aligned with the purpose of grouping within a GRACIOUS IATA.

3.3.3.1 Instructions to use an IATA

The IATA structures can be viewed as addressing three different degrees of detail:

- **IATA** – These activities apply across the whole IATA and are generally done at the entry to and exit from the IATA.
- **Decision Node** – Although the same DN might appear in several IATAs, the user should be aware that the structure and the outcome of a DN may vary (e.g. Does the DN lead to another DN or a final grouping conclusion?). In addition, the purpose of grouping will impact on how a DN is addressed.
- **Tier** – An individual tier will consist of a small number of specific methods. As the GRACIOUS Framework covers all aspects of risk assessment of NFs, the recommendations given here are general and will need to be adapted to the characteristics of the individual study.

In order to assist the user in distinguishing these different levels of recommended activities, they will be colour-coded in line with the colors used above.

1. **Identify the IATA associated with the chosen pre-defined hypothesis**

Once the user of the Framework has identified a pre-defined hypothesis that they feel is appropriate to the purpose and the candidate NFs, the Framework will guide them to the IATA that is associated with the pre-defined hypothesis. The IATA describes the strategy developed by the GRACIOUS project that is used to provide the scientific justification for a final grouping conclusion. The IATA associated with each pre-defined hypothesis can be found in the GRACIOUS project deliverables (https://www.h2020gracious.eu/).

It is recommended that the user familiarises themselves with the IATA and the DNs therein. The user will need to use their own experience and knowledge of the purpose for grouping to assess whether they may need to proceed through higher tiers at each DN or whether their purpose allows them to use lower tier studies only. The user must be aware that they will need to be flexible once they start the IATA, so that if the results of a tier 1 study do not allow a justified conclusion to be drawn, they will proceed to tier 2 or tier 3 even if this was not intended at the start of the IATA.

2. Identify all IATA data requirements and construct the data matrix

A data matrix is used to collate all the results from each study associated with an IATA. Each pre-defined hypothesis will have its own unique data matrix.

3. Populate the data matrix with available information and identify data gaps

Before starting the first DN in the IATA, the data matrix can be populated with available existing data, such as:

- Basic Information – Every data matrix will contain the physicochemical information requirements of the Basic Information step of the Framework for each of the NFs in the assessment.
- Information generated when identifying a pre-defined hypothesis – As described in the previous section, some information beyond the Basic Information may be needed to identify a single pre-defined hypothesis from a shortlist.
- Available information – Research projects have generated a huge amount of data on various nanoforms, much of which is readily accessible from databases such as eNanoMapper (see section 3.3.3.4).

Examination of the section of the data matrix relevant to each DN will identify any data gaps that need to be filled by new studies.

4. Identify potential *source*(s) for read-across (if required for the purpose)

The purpose of the grouping exercise should be used to identify the factors or limitations that will define a suitable *source*. For example, if the purpose is to provide the data needed for a specific hazard endpoint required by REACH, the *source* would need to be the same substance as the candidate NFs. If the purpose is SbD there would be more freedom to use a *source* that was not the same substance as the candidates.

Identification of potential *source*(s) can be done at any time in the Framework,

a) When identifying candidate NFs in the Basic Information step of the Framework

The user may already be aware of a (group of) NF(s) in their portfolio with the relevant data, so it might be better to identify the potential *source* NF with the required hazard data at this point. This approach ensures the Basic Information is collected at the same time as the other candidate NFs.

b) After collating available information

It is at this point in the Framework that the user can do a full data gap assessment. If no candidate NF is a potential *source*, the user may decide to add a NF (or non-NF) to their candidates to act as a source for read-across. Such an additional NF (or non-NF) may come from an external data-source such as the eNanoMapper database. The user must be aware that one of the principal rules of the GRACIOUS Framework is that DNs and the tiers therein should, where possible, be addressed using the same method for each candidate. When choosing to use a NF from an external database as a potential *source*, the user must ensure it is either available for testing or that it has the data required to complete the data matrix for the IATA.

c) During the use of the IATA

As the user moves through the DNs of the IATA, the potential *source* may be discounted from the group or a potential *source* was not identified. In this case, the user may decide to assess the grouping hypothesis before any chronic *in vivo* studies are commissioned. Based on the outcome of the similarity assessment, the user could choose a single NF in the group to use when commissioning the *in vivo* study and then use this as the *source* from which to read-across to all other group members.

The user must decide on the best approach for their particular situation but must be aware that results could force them to change their approach during the use of the GRACIOUS Framework. The user can now enter the IATA and start the first DN.

5. Examine the data requirements of the DN, including the options to move on from the DN

Each DN is intended to answer a specific question that is key to reaching a scientifically justified grouping conclusion. The DN may be split into three tiers that are differentiated by increasingly complex data requirements. The user needs to examine each tier of the DN with their purpose for grouping in mind to understand precisely what information is generated at each tier and the limitations or caveats of using this information to make a grouping decision.

6. Perform study(ies) within tier 1 of DN1 in the IATA

Some basic aspects need to be addressed for all studies.

- Consistency of method across all candidate NFs – A justified comparison of NFs can only be made if the same method is used across all NFs. Particular care needs to be taken if some of the data are extracted from databases.
- Representative test materials (RTM) should be included in each assay as internal controls for assay performance and to serve as a point of reference to support the interpretation and assessment of results obtained on a new test material.
- What method will be used to decide on the next step? Some DNs will require NFs to meet a quantitative threshold value and others will require a comparison between the candidate NFs and/or the RTM (see Section 3.3.5).
- Quality (see Section 3.3.4).
- Confidence in results – Even if a high-quality study is performed, there may be a large standard deviation seen in the final results that makes drawing a conclusion difficult, especially if this deviation falls across a threshold stipulated in the DN. Understanding this standard deviation is important for deciding whether the tier 1 results are sufficiently conclusive to move to the next DN or whether tier 2 studies are required.

This guidance document cannot give specific guidance for each different study, so the user should consult the articles referenced at the end of this section for specific support.

Can the bullet points above be tidied up?

7. **Decide whether DN1 is sufficiently addressed by tier 1 or whether higher tier studies are needed, then perform higher tier studies if required.**

Every DN will have different criteria for deciding whether sufficient information exists to make a decision or whether higher tier studies are needed. The decision will also be made on the basis of a combination of the factors discussed previously, i.e. purpose of grouping, quality and confidence in results. For examples of how these decisions have been made in "real-life" examples, please read the research published by the GRACIOUS project detailed in section 3.3.3.4. The user should progress from DN to DN through the IATA until a conclusion on grouping can be scientifically justified.

8. **Repeat stages 6 and 7 for each DN until all DNs have been addressed**

It is possible that the conclusion from a DN will be that none of the NFs can be grouped under the hypothesis. In this case the user can leave the IATA without moving to subsequent DNs. They will then need to decide whether to select a different grouping hypothesis or to conclude that the candidate NFs cannot be grouped. To accept a grouping hypothesis for some or all of the candidate NFs, all DNs in an IATA should be addressed to reach a final grouping decision.

9. **Use the results from all the DNs to identify a grouping conclusion**

Once sufficient data has been collated in the data matrix to make a scientifically justified conclusion, a group of NFs can be defined according to the original purpose of the exercise. The GRACIOUS Framework recommends that a quantitative measure of similarity of NFs in a group is used as a key tool in the grouping justification (Section 3.3.5).

The actions that can be taken if a group is identified are covered in section 3.4. Where some or all the candidate NFs cannot be grouped, the user of the Framework will need to decide on their next step based on the original purpose of the grouping. Some possible examples are given below, but this is not an exhaustive list.

a) If the purpose of the grouping was to define the boundaries of a group to support scientific understanding, the ungrouped NFs have served their purpose and no further work is needed.
b) Other pre-defined hypotheses may be more appropriate for the ungrouped NFs. The user should use the information in the original data matrix to identify another pre-defined hypothesis for potential grouping. If an alternative pre-defined hypothesis is identified, the user should be aware that they might be able to transfer data from the original data matrix to the data matrix of the new pre-defined hypothesis.
c) Consider adapting the original pre-defined hypothesis to reflect observations made during original experiments.

3.3.3.2 Worked Example 1: Use of HARN IATA for grouping MWCNTs for regulatory purposes

This section covers the bulk of experimental work and includes many points where the user needs to make their own decisions on how to interpret results and how to proceed through the Framework. Two worked examples of how GRACIOUS project partners have applied an IATA for a specific pre-defined hypothesis

are presented here. These examples highlight the considerations that need to be made and the obstacles that were met. The user of the Framework needs to be aware that their situation is unlikely to be identical to these worked examples, even if they are using the same pre-defined hypothesis. Publications detailing other pre-defined hypotheses and case studies are referenced at the end of the section.

This section provides an example for human toxicity. An example for environmental toxicity is provided below (section 3.3.3.3). If the user wants to know more about the development of the hypothesis and IATA used in this worked example on human toxicity, please read Murphy *et al*. (2021).

Context

In this worked example, the '**User'** is a MWCNT manufacturer with a panel of MWCNTs that require REACH registration but currently have differing degrees of hazard data available.

Their **purpose** is to support a read-across of available hazard data between group members for regulatory hazard assessment.

The context for this grouping exercise is the occupational exposure to MWCNTs during primary production and packaging, as well as incorporation of MWCNTs into nano-enabled products. Potential aerosolization of MWCNTs during the production processes indicates inhalation as the route of primary concern. The occupational setting also suggests the potential for repeated exposure. Therefore, a regulatory hazard endpoint of primary interest is repeated dose toxicity by inhalation, specifically whether it is possible to use one OECD TG 412 (28-day subacute inhalation toxicity) and/or OECD TG 413 (90-day subchronic inhalation toxicity) study to cover all candidate NFs.

Basic Information Step leading to identification of pre-defined hypotheses

This has been detailed in section 3.3.1.2 of this document, but the key outcomes are shown in table 3.3.3.

Table 3.3.3: Basic Information for 5 NFs of MWCNT

NF	Carbon (%)	Length Mean± SD (µm)	Diameter Mean± SD (nm) (range)	Shape* Aspect Ratio (D3:D1)	BET (m^2/g)	Level of Existing Hazard Data
MWCNT-A	86.2	0.85±0.10	11±3 (6-17)	Elongated 77.27	254	None
MWCNT-B	99.7	4.0±0.37	67±24 (24-138)	Elongated 59.7	18	Acute in vitro
MWCNT-C	96.1	1.4±0.19	11±3 (7-20)	Elongated 127.27	226	None
MWCNT-D	99.1	0.4±0.03	12 ±7 (5-37)	Elongated 33.33	135	Acute in vitro STIS
MWCNT-E	99.6	5.7±0.49	74 (29-173)	Elongated 77.02	26	Acute in vitro OECD TG 413

H-I-1: "*Respirable, bio persistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of HARNs can occur resulting in lung toxicity*."

It was noted that **H-I-1** and **H-I-2** are related but that they relate to different endpoints (lung toxicity (**H-I-1**) or mesothelioma (**H-I-2**)). The purpose of the grouping is to provide the data required for inhalation toxicity endpoints in REACH. REACH requires registrants that put > 1000 tons per year on the market to make an assessment of carcinogenic potential, and further investigation for lower tonnages may be

required if the data supports this. **H-I-1** is used in this example, but the user should be prepared to extend the work to **H-I-2** if the data supports this.

This worked example now discusses the application of the instructions provided in section 3.3.3.1.

1. Identify the IATA associated with the chosen pre-defined hypothesis

The IATA associated with **H-I-1** is shown in Figure 3.3.6.

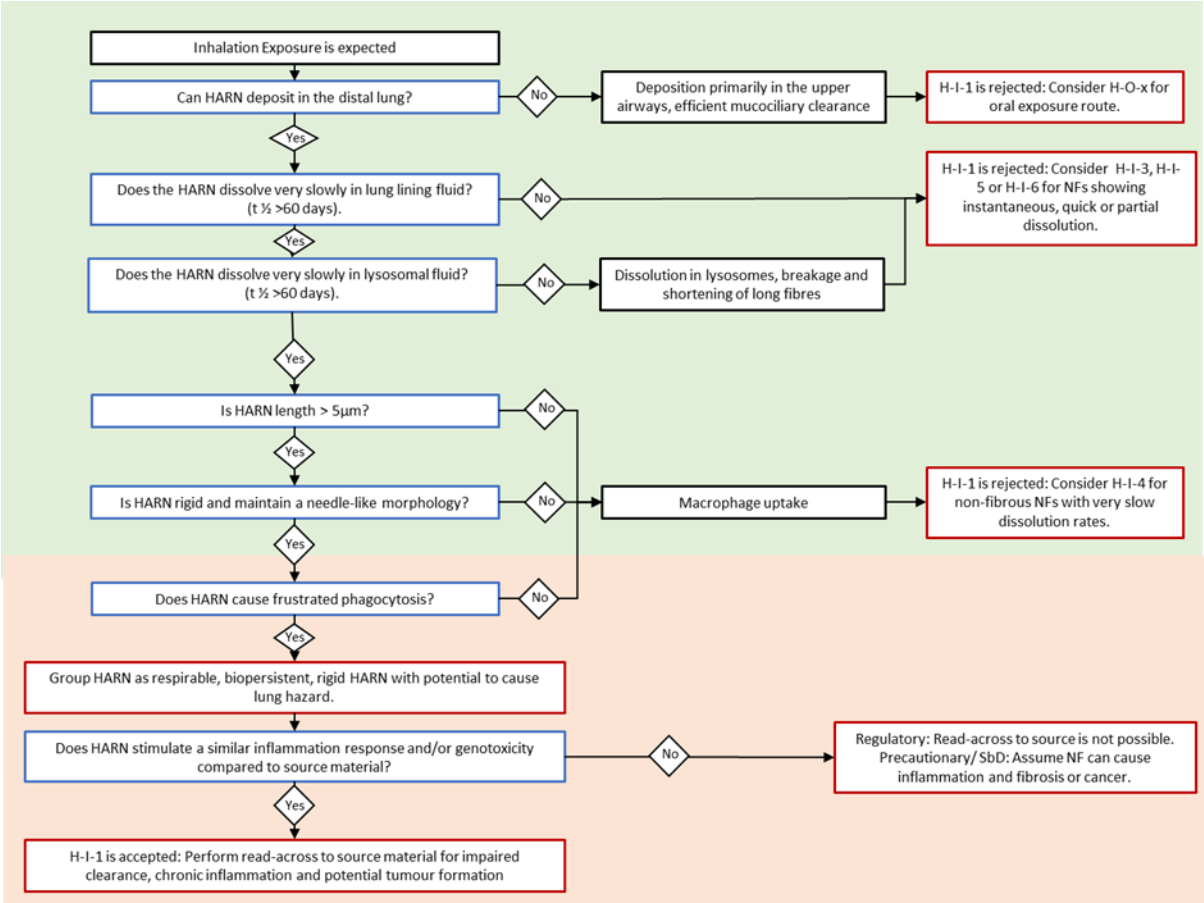


Figure 3.3.6: IATA for HARN hypothesis, H-I-1. Blue boxes indicate IATA decision nodes, red boxes indicate the outcome of the decision nodes, black boxes provide explanatory information relevant for the interpretation of the decision node outcome.

This IATA builds on the well-characterised structure activity relationship (SAR) which predicts the pathogenic potential of mineral fibres. The DNs are designed to identify candidate HARNs which meet the criteria for fibre pathogenicity, which forms the basis of the grouping hypothesis, and reject HARNs which do not. A similarity assessment across all DNs can be used to strengthen a grouping decision and support the possibility to use read-across approaches to fill data gaps for pulmonary hazard endpoints.

HARNs that have been eliminated from the group, through the IATA process, may be assessed for the applicability of different pre-defined hypotheses which examine alternative hazard mechanisms.

2. Identify all IATA data requirements and construct the data matrix

The individual studies required for each tier in each DN are shown in Table 3.3.4.

*Table 3.3.4: The studies required to address each tier at each decision node (DN) of the pre-defined hypothesis **H-I-1**.*

DN1: Can HARN deposit in the distal lung?	DN2: Does the HARN dissolve very slowly in lung lining fluid?	DN3: Does the HARN dissolve very slowly in lysosomal fluid?	DN4: Is HARN length >5µm?	DN5: Is the HARN rigid and maintain fibrous, needle-like morphology?	DN6: Does the HARN cause frustrated phagocytosis?	DN7: Does HARNs stimulate a similar inflammation response and/or genotoxicity to source material?
Tier 1						
Review existing data sets						
Estimation of D_{ae} from HARN size measurements by TEM/SEM and density measurement	Batch dissolution test in lung lining fluid (pH 7.4) or Dissolution in continuous flow system in lung lining fluid (pH 7.4)	Batch dissolution test in lysosomal fluid (pH 4.5) or Dissolution in continuous flow system in lysosomal fluid (pH4.5)	HARN size measurements by TEM/SEM	Measure diameter of HARN by TEM	Inflammasome activation: • IL-1β release • CathepsinB activity and /or release • Lysosomal Disruption	Inflammation potency: *in vitro* testing using cell lines Acute Endpoints: • Cytotoxicty • Cytokine release • Oxidative Stress • DNA damage
Tier 2						
Review existing data sets						
Measurement of MMAD by cascade impactor from an airborne dispersion of the material Lung deposition modelling: • Multiple Particle Path Deposition Model		Durability in cellular systems	HARN size measurements by TEM/SEM from an airborne dispersion of the material	HARN size measurements by TEM/SEM from an airborne dispersion of the material	*In vitro* granuloma formation	*In vitro* incubation with co-culture models of macrophages and mesothelial cells or 3D microtissue models Acute Endpoints: • Cytokine release • DNA damage Chronic: • Granuloma formation • Cell transformation
Tier 3						
Review existing data sets						
Quantification of lung burden after in vivo inhalation studies (OECD TG 412/413). Initial timepoint to measure deposition in distal lung Longer timepoint to measure bio persistence					Intraperitoneal/ Intrapleural instillation: Acute Endpoint: • Inflammation, • Oxidative DNA damage Chronic: • Fibrotic lesion • Mesothelioma	

By combining the data requirements from both the Basic Information and the IATA, a blank data matrix can be constructed (Table 3.3.5).

*Table 3.3.5: A blank simplified data matrix for IATA **H-I-1**.*

Decision Node	Tier	Study	NF1	NF2	NF3…
Basic Information		Carbon %			
		Length			
		Diameter			
		Shape Aspect Ratio			
		Specific Surface Area			
DN1	1	Mean Diameter			
		Density			
	2	MMAD			
	3	Quantification of lung burden			
DN2	1	Dissolution rate in lung lining fluid			
	2	Dissolution rate in intracellular environment			
DN3	1	Dissolution rate in lysosomal fluid			
	2	Dissolution rate in intracellular environment			
DN4	1	Average fibre length			
		Fibre length distribution			
	2	Fibre length distribution in airborne dispersion			
DN5	1	Fibre width distribution			
		Agglomeration state			
	2	Fibre width dispersion in airborne dispersion			
DN6	1	IL-1β release			
		CathepsinB activity/release			
		Qualitative assessment of protrusion/piercing cell membrane			
		Qualitative assessment of lysosomal disruption			
	2	Development of 3D macrophage granulomas			
	3	*In vivo* hazard response			
DN7	1	*In vitro* testing using cell lines Cytotoxicty Cytokine release Oxidative Stress DNA damage			
	2	Cytokine release DNA damage Granuloma formation Cell transformation			
	3	*In vivo* hazard response			

3. Populate the data matrix with available information and identify data gaps

The available data was taken from the Basic Information and any other relevant data that had been previously generated (Table 3.3.6).

*Table 3.3.6: Available information for the worked example of **H-I-1** before entering the IATA.*

Decision Node	Tier	Study	MWCNT-A	MWCNT-B	MWCNT-C	MWCNT-D	MWCNT-E
Basic Information		Carbon %	86.2	99.7	96.1	99.1	99.6
		Length	0.85±0.10	4.0±0.37	1.4±0.19	0.4±0.03	5.7±0.49
		Diameter	11±3 (6-17	67±24 (24-138)	11±3 (7-20)	12 ±7 (5-37)	74 (29-173)
		Shape Aspect Ratio	Elongated 77.27	Elongated 59.7	Elongated 127.27	Elongated 33.33	Elongated 77.02
		Specific Surface Area	254	18	226	135	26
DN1	1	Mean Diameter	11±3 (6-17	67±24 (24-138)	11±3 (7-20)	12 ±7 (5-37)	74 (29-173)
		Density					
	2	MMAD					
	3	Quantification of lung burden					
DN2	1	Dissolution rate in lung lining fluid					
	2	Dissolution rate in intracellular environment					
DN3	1	Dissolution rate in lysosomal fluid					
	2	Dissolution rate in intracellular environment					
DN4	1	Average fibre length					
		Fibre length distribution					
	2	Fibre length distribution in airborne dispersion					
DN5	1	Fibre width distribution	11±3 (6-17	67±24 (24-138)	11±3 (7-20)	12 ±7 (5-37)	74 (29-173)
		Agglomeration state					
	2	Fibre width dispersion in airborne dispersion					
DN6		IL-1β release					
		CathepsinB activity/release					
	1	Qualitative assessment of protrusion/piercing cell membrane					
		Qualitative assessment of lysosomal disruption					
	2	Development of 3D macrophage granulomas					
	3	*In vivo* hazard response		Data from acute study		Data from acute study	OECD TG 413
DN7	1	*In vitro* testing using cell lines Cytotoxicity Cytokine release Oxidative Stress DNA damage					
	2	Cytokine release DNA damage Granuloma formation Cell transformation					
	3	*In vivo* hazard response					

In this worked example, some Basic Information can be used to satisfy some of the information requirements of tier 1 testing. The only NF that has OECD TG 413 data available is MWCNT-E. All other studies in the matrix have data gaps. The user should be aware that not every data gap needs to be filled,

as the unique properties of the substance under assessment might mean it is not scientifically justified to commission some studies. Also, the lower tier results may already sufficiently justify moving on to the next DN or drawing a grouping conclusion without needing to do the higher tier study.

4. Identify potential source(s) for read-across (if required for the purpose)

The purpose of this grouping is to provide information on repeated dose toxicity for all NFs. As MWCNT-E has data available from an OECD TG 413 study, it is considered as the potential *source* NF at this point in the Framework.

5. Examine the data requirements of each DN, including the options to move on from the DN

Examination of the exact data requirements of each DN and each tier may reveal issues that will make performing the studies or drawing useful conclusions difficult due to the physicochemical properties of the candidate NFs under investigation. It may also identify opportunities to run some studies in parallel or satisfy multiple data requirements with a single study. Examples from this worked example include:

- **DN1:** Can HARNs deposit in the distal lung? - **Tier 1**: Estimation of Aerodynamic radius (D_{ae}) derived from median diameter and density

The potential wide variation in fibre morphology and agglomeration status in a HARN sample may lead to a high level of uncertainty in the DN outcome when based on a Tier 1 estimation of D_{ae} (Figure 3.3.7).

Figure 3.3.7: Increasing heterogeneity within the sample leads to increasing uncertainty in estimation of D_{ae}. For 'Ideal' homogonous straight fibres, tier 1 is sufficient to make a prediction with a high level of confidence based on *in silico* modelling from estimated D_{ae}. The more the sample deviates from the ideal, the more the uncertainty increases. tier 2 measurement of aerodynamic diameter may be required to confirm potential for deposition in the distal lung and comparison of modelled deposition fractions.

Qualitative electron microscopy (EM-Basic PC Information) can provide an indication of the heterogeneity within the HARN sample, i.e. the presence of fibrous and particulate fractions and level of agglomeration. The level of heterogeneity and user needs will determine whether a tier 1 estimation of D_{ae} will prove sufficient to support a conclusion on the likelihood of deposition in the distal lung. If a high level of heterogeneity is identified in the estimation of D_{ae} for HARNs such as MWCNTs the predicted D_{ae} is not considered appropriate for similarity assessment between the MWCNTs. Therefore, conducting a similarity assessment between the MWCNTs to support read-across requires measured D_{ae} from tier 2 assays.

- **DNs 2 and 3:** Do the HARNs dissolve very slowly in lung lining fluid/lysosomal fluid? **Tier 1**: Batch dissolution in continuous flow system in fluid

DNs 2 and 3 could be completed in parallel at tier 1, as they utilise the same model set-up, the only change in condition being the different incubation media. Both DNs are required to complete the IATA. In the case of carbon allotropes such as MWCNTs, it is difficult to define or measure dissolution in any media. Expert judgement can justify that they can be assumed to dissolve very slowly.

6. **Perform study(ies) within tier 1 of DN1 in the IATA**
7. **Decide whether DN1 is sufficiently addressed by tier 1 or whether higher tier studies are needed, then perform higher tier studies if required.**
8. **Repeat stages 6 and 7 for each relevant tier for each DN until all DNs have been addressed**

This worked example takes the user through all the studies in each DN, showing how the results of the studies in the IATA have been used to decide how to progress through the tiers in a DN and how to complete the data matrix. Therefore stages 5, 6 and 7 are all addressed concurrently.

Full details of the experimental work and decision-making process can be seen in Murphy et al. (2021). A summary of each DN is given here.

DN 1: Can HARNs deposit in the distal lung?

Tier 1: Estimation of D_{ae} derived from median diameter and density

Although the data is available from the Basic Information, heterogeneity of the candidate NFs means that a useful conclusion cannot be made.

Tier 2: Measurement of Mass Median Aerodynamic Diameter (MMAD)

Studies on all candidate NFs displayed a $D_{ae} < 2$ µm (Table 3.12), meaning that deposition in the distal lung is likely and progress to DN2 without tier 3 is appropriate.

Tier 3: Quantification of Lung Burden after inhalation exposure

Not required as tier 2 information is sufficient.

<u>Conclusion from DN</u>

Since particle deposition in the distal lung is likely for all NFs, all candidates proceed to DN2.

DN2 and 3: Do the HARNs dissolve very slowly in lung lining fluid/lysosomal fluid?

The DN are run in parallel (see above).

Tier 1: Batch dissolution in artificial lung fluids.

As stated above, dissolution studies on MWCNTs are unnecessary because all carbon allotropes have previously been shown to dissolve very slowly or to be completely insoluble. Move to DN4.

Tiers 2 and 3 – Unnecessary based on tier 1 decision.

Conclusion from DN

All candidates proceed to DN4 due to low or lack of dissolution.

Decision Node 4: Is the length of HARNs > 5µm?

Tier 1: HARN size measurements by TEM/SEM from a suspension of the material

A summary metric reporting median or mean fibre length of < 5 µm is not sufficient to meet the criteria for this DN due to the potential of wide size distributions within a sample, as demonstrated in Table 3.13. Rather the proportion of fibres > 5µm should be reported. A pragmatic threshold of 10 % has been set to differentiate samples composed of very few long fibres. Such low proportions of very long fibres are considered to represent a very low hazard.

Tier 2: HARN size measurements by TEM/SEM from an airborne dispersion of the material

It is worth noting tier 1 assessment of the size profiles of HARNs in suspension may not necessarily reflect the size profile of the HARN when aerosolized. Tier 2 requires the measurement of shape and size profiles from aerosolized samples and will provide confirmation that the HARNs in aerosolized form meet the threshold. The threshold to satisfy tier 2 is set according to WHO criteria of 0.1 % of particles having length > 5 µm (and an aspect ratio of > 3:1), with both individual fibres and fibrous agglomerates included (World Health Organisation, 1996). For many examples where the purpose is to provide the data required for REACH regulatory endpoints, much of this data should be already available in the Dustiness endpoint that is mandatory for all forms, including NFs.

Table 3.3.7: Data measured on the five candidate nanoforms for DN4, tiers 1 and 2 of IATA H-I-1.

NF	Tier 1		Tier 2
	Mean length (µm) ± SD	Range (µm)	% fibres > 5µm
MWCNT-A	0.85±0.10	0.1 - 10	6
MWCNT-B	4.0±0.37	3 - 20	33
MWCNT-C	1.4±0.19	0.2 - 10	8
MWCNT-D	0.4±0.03	0.05 - 6	1
MWCNT-E	5.7±0.49	5 - 40	40

Conclusion from DN

All candidates proceed to DN5.

Decision Node 5: Is the HARN rigid and does it maintain a fibrous, needle-like morphology?

The bending rigidity of a HARN determines whether the HARN will maintain their confirmation as fibres during handling, abiotic dissolution, or cell interactions. Diameter, alongside the elastic modulus (Young's modulus) of a material can be used to predict fibre rigidity from the Euler Buckling theory at an individual fibre level. At tier 1, the IATA fibre diameter is therefore used as an indirect indicator of rigidity of an individual HARN. A threshold of 30 nm has been set based on evidence which suggests that this is a critical threshold for fibre buckling for a number of relevant materials under compressive forces of biological process such as phagocytosis (~10^{-19} Nm^2).

Tier 1: Measure diameter of HARN by TEM and make a qualitative assessment of agglomeration state

High power/high magnification EM images are used to measure the HARN diameter on a constituent particle level to address the DN threshold, i.e. a diameter > 30 nm indicates that the HARN is rigid.

It is important to note that although a HARN may not be considered rigid at a single fibre level, the formation of agglomerates of multiple fibres may lead to the formation of aligned bundles which present a rigid, fibrous, needle-like morphology (Figure 3.3.8). To account for this phenomenon, low power/low magnification EM representative of the agglomeration state/assembly structure of the HARN should also be provided.

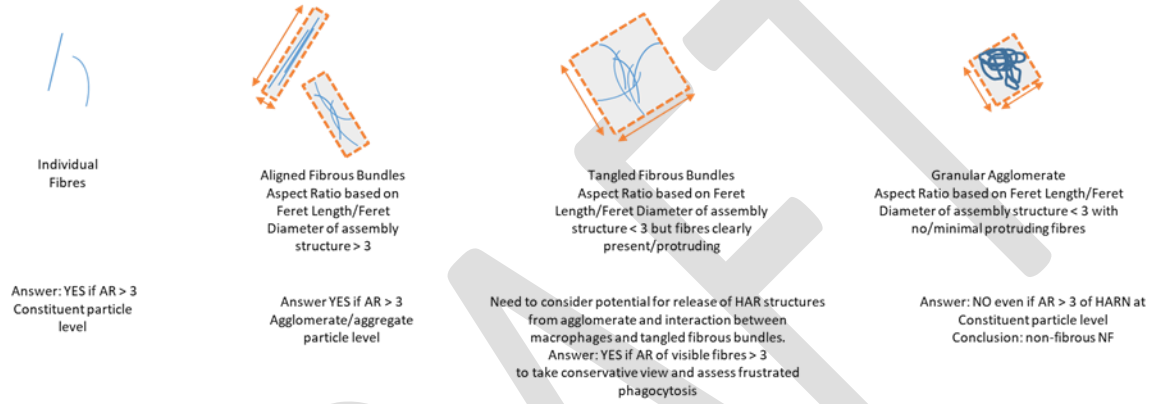


Figure 3.3.8: Diagrams showing a quantitative assessment of the aspect ratio of possible agglomerate states of fibrous nanoforms such as MWCNTs.

The results used in this worked example are shown in Table 3.3.8.

*Table 3.3.8: Measured fibre diameter and agglomerate state required for DN5, tier 1 of IATA **H-I-1***

NF	Is the HARN rigid and does it maintain a fibrous, needle-like morphology?
MWCNT-A	Fibre diameter < 30 nm Agglomeration state: Tangled agglomerates which appear more granular than fibrous are predominant in the supporting SEM.
MWCNT-B	Fibre diameter > 30 nm Agglomeration state: Rigid fibre-like structure is supported by SEM.
MWCNT-C	Fibre diameter < 30 nm. Agglomeration state: Tangled agglomerates which appear more granular than fibrous are predominant in the supporting SEM.
MWCNT-D	Fibre diameter < 30 nm. Agglomeration state: Tangled agglomerates which appear more granular than fibrous are predominant in the supporting SEM.
MWCNT-E	Fibre diameter > 30 nm Agglomeration state: Rigid fibre-like structure is supported by SEM.

Mean Diameter reported for Basic Information suggests that only two MWCNT test samples meet the 30 nm threshold; **MWCNT-B** and **MWCNT-E**, whose rigid fibre-like structure is supported by SEM. These MWCNTs can be considered a provisional group. It was decided that this was sufficient data to move these NFs to DN6.

MWCNT-A, **MWCNT-C** and **MWCNT-D** do not meet the threshold. Tangled agglomerates which appear more granular than fibrous are predominant in the supporting SEM. These samples were removed from the group of candidate NFs for the HARN hypothesis. However, as the purpose of the work was to provide the data needed for a REACH inhalation toxicity endpoint, these forms still need relevant endpoint data. Therefore, they were considered for inclusion in one of the other inhalation hypotheses. As existing data generated during testing for **H-I-1** can be used in the other inhalation hypotheses, the three MWCNTs have been shown to be respirable and have a very slow dissolution rate, so they will be included as candidate NFs in **H-I-4**, i.e. "Respirable NFs with a very slow dissolution rate: Following chronic inhalation exposure, accumulation of NFs in the lungs can lead to long-term toxicity" (Figure 15).

Conclusion from DN

MWCNT-B and **MWCNT-E** proceed to DN6 for the IATA associated with hypothesis **H-I-1**. The hypothesis **H-I-1** does not apply to **MWCNT-A**, **MWCNT-C** and **MWCNT-D**, so they cannot be grouped based on this hypothesis. Grouping based on hypothesis **H-I-4** should be considered.

Decision Node 6: Does the HARN cause frustrated phagocytosis?

Frustrated phagocytosis is a direct outcome of the inability of macrophages to completely engulf and passivate high aspect ratio materials, leading to the activation of NALP3 inflammasome, via lysosomal disruption, and resulting in the release of the pro-inflammatory cytokines, IL-1β and IL-18 (Murphy *et al.*, 2012; Palomäki *et al*., 2011). This DN indicates a biological hazard outcome linked to fibre morphology. Frustrated phagocytosis is however difficult to define, and no standard method to assess frustrated phagocytosis exists yet.

Tier 1: Incubation of HARN with macrophages in submerged cell culture.

A panel of quantitative and qualitative endpoints indicative of frustrated phagocytosis (Table 3.3.9) should be assessed to address this DN (see Murphy *et al.* (2021)) for justification of selection of endpoint panel) and both biochemical read-outs as well as supporting microscopy images should be included.

Table 3.3.9: Experimental options for investigating the potential for frustrated phagocytosis in DN6, tier 1 of IATA H-I-1.

Quantitative Endpoints	Qualitative Endpoints
IL-1β release: Quantitative measure of mature IL-1β release into cell culture supernatant after short-term incubation of HARN with macrophages, indicative of NALP3 inflammasome activation. PATROLS SOP- THP-1 human monocytic/macrophage cell line, 24h NF exposure. **IL-18 release**: Quantitative measure of mature IL-1β release into cell culture supernatant after short-term incubation of HARN with macrophages, indicative of NALP3 inflammasome activation. PATROLS SOP- THP-1 human monocytic/macrophage cell line, 24h NF exposure. **Cathepsin B extracellular release/activity**: Measurement of CathepsinB enzymatic activity in cell culture supernatant or quantitative measure of CathepsinB release into cell culture supernatant. Although not direct measure can be indicative of potential of inflammasome activation.	**Protrusion/piercing cell membrane**: Images which show failure of macrophages to completely engulf HARN with external cell membrane by light microscopy and/or SEM/TEM or failure to form a phagolysosome which encapsulates the HARN as visualized by TEM can be used in support of quantitative, biochemical endpoints to strength DN outcome and better illuminate/confirm proposed MoA. **Lysosomal Disruption**: Fluorescent microscopy analysis of lysosomal disruption using lysosome specific fluorescent dyes such as Lysotracker, Magic Red, Acridine Orange. Qualitative assessment based on visualisation of loss of punctate staining indicative of intact lysosomes and diffusion of dye throughout the cytoplasm.

Due to the system-dependent methods to assess interactions between HARNs and cells, and the inherent biological variability in these assays, it is not possible to define a threshold whereby a HARN can be characterized as causing frustrated phagocytosis to a pathologically-relevant degree. Therefore, answering the DN requires a comparison of the similarity of responses between the HARNs under investigation for each endpoint (between each other to support grouping), alongside a comparison of potency of HARN responses to well-characterized negative and positive RTMs which have been demonstrated to elicit the endpoint response under investigation.

Table 3.3.10: Results of in vitro studies investigating frustrated phagocytosis in DN6, tier 2 of IATA H-I-1 for the two remaining candidates for grouping.

NF	Does HARN cause frustrated phagocytosis?
MWCNT-B	IL-1β release measured after 24 hour exposure to THP-1 macrophages shown to generate significant increases in IL-1b release over control, in line with positive benchmark material
MWCNT-E	IL-1β release measured after 24 hour exposure to THP-1 macrophages shown to generate significant increases in IL-1b release over control, in line with positive benchmark material

Both remaining candidate NFs have a qualitatively similar response in terms of inducing frustrated phagocytosis (Table 3.3.10).

From the above DNs it is clear that only two of the candidate HARNs can be considered to form a group defined by the hypothesis. To further strengthen the argument to read-across hazard data between group members, the level of similarity between the HARNs should be evaluated across all DNs of the IATA. A robust read-across argument will require justification of the level of (dis)similarity considered to be acceptable, i.e. not sufficient to drive an alternative hazard outcome.

Conclusion from DN

MWCNT-B and MWCNT-E demonstrate a similar ability to cause frustrated phagocytosis. Both remaining candidates proceed to DN7.

DN 7: Does the HARN elicit a similar inflammatory and/or genotoxicity response to the source material?

As the identified potential source, **MWCNT-E**, is still a candidate NF: the grouping to allow read-across from this NF to all other group members (**MWCNT-B**) is still possible. To further strengthen the mechanistic underpinning of the group, further assessment of the similarity in hazard outcomes between group members may be carried out, and data from simple tier 1 *in vitro* assays, more complex *in vitro* and short-term *in vivo* assays predictive of long-term hazard outcomes addressed in the OECD TG 413 may be incorporated. These studies have not been performed at the time of writing this Guidance Document.

9. Use the results from all the DNs to identify a grouping conclusion

To draw a final grouping conclusion, the similarity between the source, **MWCNT-E**, and the targets, **MWCNT-B** only, should be quantitatively measured using the approaches suggested in section 3.3.5 after assessment of the quality of the data set as described in section 3.3.4.

3.3.3.3 Worked example 2: Use of the general aquatic IATA to identify groups where either NF particles, solutes or a combination of both drive lethal and sub-lethal toxicity to representative aquatic species

Context

This is a worked example for hypothetical candidate NFs, in which the user is a manufacturer of several NFs of silver (< 100 t/year) for which some intrinsic physicochemical data exists. Releases to the aquatic environment are expected through the use phase of the product life cycle. It is also suspected that the NFs will dissolve to some extent in water on the basis of their chemical identity. Silver nitrate ($AgNO_3$), the soluble salt of the metal has been identified as a potential *source* non-NF for read-across of aquatic toxicity endpoints for which *in vivo* toxicity data is already available. The **purpose is regulatory,** to see if the new candidate NFs may be grouped with the existing source solutes of the material to reduce *in vivo* testing requirements for aquatic toxicity in REACH.

The regulatory hazard endpoint of primary interest in this worked example is aquatic pelagic toxicity as part of the information requirements in the hazard assessment for REACH (REACH Annex VIII). The case study investigates whether it is possible to use existing short-term toxicity testing on fish (OECD TG 203) for the solute of the nanomaterial ($AgNO_3$) as a source to read-across to all candidate NFs.

Multiple grouping outcomes are possible under the generic aquatic IATA:

- Fate and toxicity of the NFs are similar to the soluble source material, justifying read-across to this source.
- Read-across to the soluble source material is not justified, but NFs are sufficiently similar to allow for data from one grouped candidate NF to be suitable for read-across to other members of the group. Data generation for one NF may be required to generate a *source*.

Basic Information step leading to identification of pre-defined hypothesis

General Basic Information requirements have been detailed in the previous sections of this document, but the key outcomes for this worked example are shown in table 3.3.11. Please note that throughout this worked example, the values for the candidate NFs are for demonstrative purposes only and are purely hypothetical to demonstrate elements of the IATA.

Table 3.3.11: Basic Information available for the 5 candidate NFs.

NF	Purity (%)	Length Mean± SD (µm)	Diameter Mean± SD (nm) (range)	Shape* Aspect Ratio (D3:D1)	BET (m^2/g)	Level of Existing Hazard Data
NF1 Ag	> 99.5 (Ag)	See diameter*	2.5±1 (2-5)	Spheroidal		None
NF2 Ag	> 99.5 (Ag)	See diameter*	20±5 (10-32)	Spheroidal		None
NF3 Ag	> 99.5 (Ag)	See diameter*	50±35 (20-110)	Spheroidal		None
NF4 Ag Nanorod	> 99.5 (Ag)	0.1 ± 0.03	20 ±3 (15 - 24)	Elongated (5)		None
NF5 AgO nanowire	46.55 (Ag)	20 ± 1.2	65 ±5 (60-70)	Elongated (307.7)		None
NF6 Ag-(Ag_3PO_4)		See diameter*	55±8 (36-68)	Spheroidal		None
$AgNO_3$ (potential source material)		NA	NA	NA	NA	Acute and chronic ecotoxicity data

* NF1, NF2, NF3 and NF6 are classed as spheroidal, with an aspect ratio <3:1, so a single dimension of constituent particle diameter is sufficient to describe the size distribution of the NFs.

Although this example is not explicitly described in Sections 3.3.1 and 3.3.2, the same principles described in this section were used to identify the pre-defined hypothesis discussed in this worked example.

The pre-defined hypothesis identified is "***E-G-1**: NFs in the aqueous environment: Following aqueous exposure dissolution rate and attachment efficiency (derived from dispersion stability) are the main driving forces that determine NF fate in aqueous environments, and are sufficient as input in fate modelling of NFs. Lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics in aqueous environments of either NF particles or solutes or both*."

This worked example now discusses the application of the instruction provided in section 3.3.3.1.

Identify the IATA associated with the chosen pre-defined hypothesis

This pre-defined hypothesis comprises 6 sub-hypotheses, based on the interplay between dissolution of solutes from the NF and the dispersion stability of the particles (Table 3.3.12).

*Table 3.3.12: Sub-hypotheses under **E-G-1** considering different grouping outcomes for NFs in the aquatic environment on the basis of their dissolution and dispersion stability*

Sub-Hypothesis	Description
E-G-1a	*"NFs with a quick dissolution rate in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the solutes."*
E-G-1b	*"NFs with a very slow dissolution rate and a stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the NFs in aqueous environment."*
E-G-1c	*"NFs with a very slow dissolution rate and a partial stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the NFs remaining in aqueous environments."*
E-G-1d	*"NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (a high toxicity ratio solute : NF allows read-across to similar solutes)."*
E-G-1e	*"NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (a low toxicity ratio solute : NF allows read-across to similar NFs)."*
E-G-1f	*"NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (an intermediate toxicity ratio solute : NF limits possibilities for read-across)."*

At this step of the Framework, the candidate NFs may propagate through any of these 6 sub-hypotheses, depending on the outcomes of DNs within the IATA. In doing so, the user will identify candidate NFs that may be grouped together with the *source* material using a single sub-hypothesis. Alternatively, for those NFs which fall outside of this group, (for example a single NF falling into one of the other sub-hypotheses), the user has several options:

- Pursue specific testing of candidate NFs that were not grouped with other available candidate/*source* materials.
- Assess hazard on a case-by-case basis without the use of grouping.

As examples, we present two cases that cover all the DNs for the hypothesis, to illustrate how the answers to DNs dictate which sub-hypothesis each candidate NF propagates through.

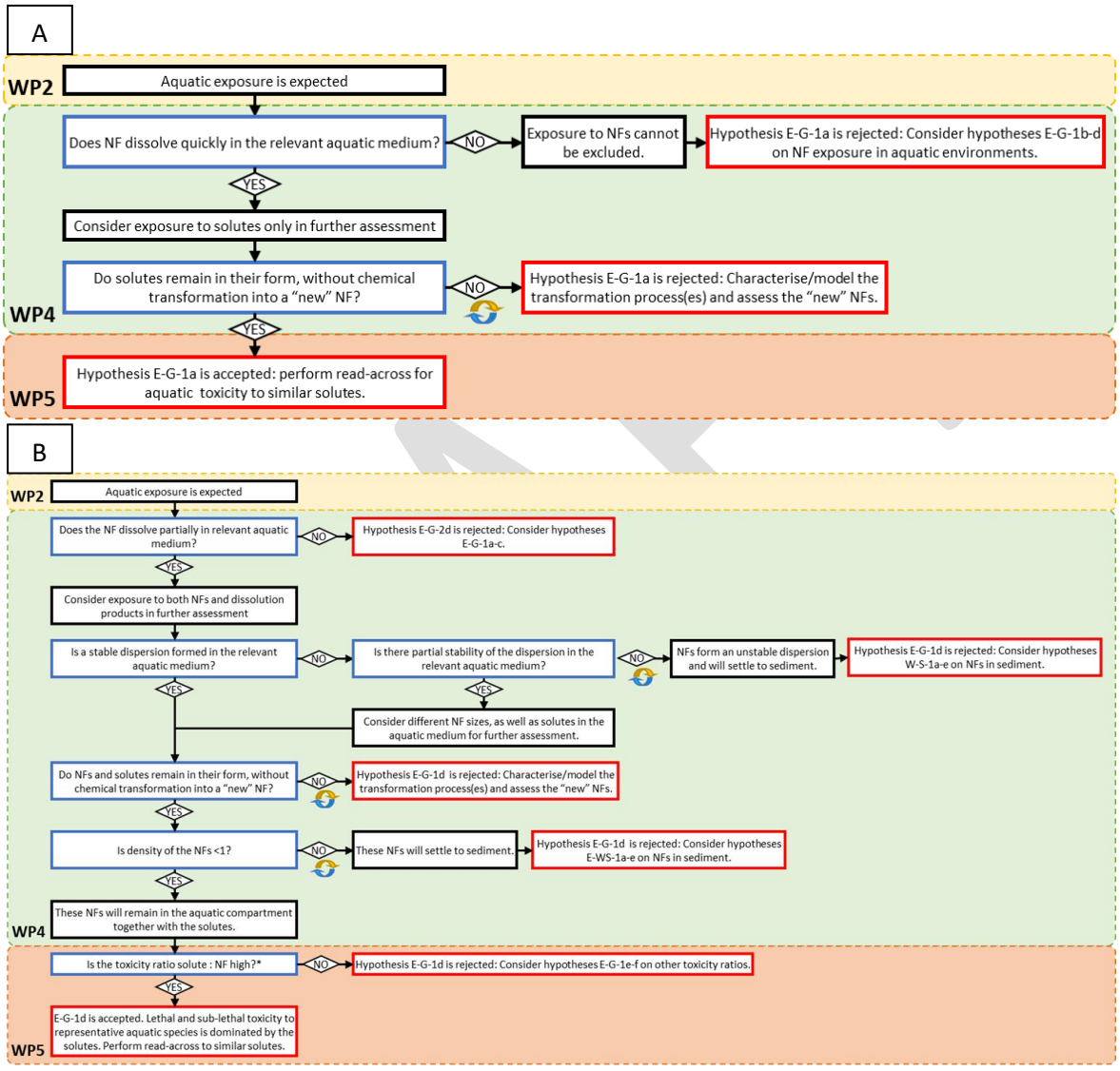


Figure 3.3.9: A) IATA for sub-hypothesis E-G-1a: NFs with a quick dissolution rate in environmentally relevant aquatic media: Following aqueous exposure, lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the solutes. B) IATA for sub-hypothesis E-G-1d: NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (a high toxicity ratio solute : NF allows read-across to similar solutes). Blue boxes indicate IATA decision nodes, red boxes indicate the outcome for the decision nodes and black boxes provide explanatory information relevant for the interpretation of the decision node outcome. The blue and yellow cycle icon represents a context switch in the tool, where a "new" candidate NF arises which may also be propagated through the IATAs. In this instance, it allows consideration of new nanoforms arising through precipitation of solutes from the original NF under environmental conditions to be taken into account.

Figure 3.3.9 presents two IATAs in which successful grouping by the sub-hypotheses justifies read-across for aquatic toxicity to similar solutes. However, the justification for reaching this read-across decision differs between sub-hypothesis **E-G-1a** and **E-G-1d**.

Figure 3.3.9A presents the IATA for sub-hypothesis **E-G-1a**, the outcome of which is that grouping supports the justification for read-across to the aquatic toxicity for solutes of the same chemical identity as the NF. In this sub-hypothesis, the candidate NF is demonstrated to dissolve at a sufficient rate prior to contact with organisms in the environment, so that only exposure to the solutes is considered in further assessment.

A similar grouping outcome is described in Figure 3.3.9B, representing the IATA for sub-hypothesis **E-G-1d**. Once again, the outcome of the group is the same as for sub-hypothesis **E-G-1a**, forming part of the justification for read-across for aquatic toxicity to similar solutes. However, the rationale behind this justification differs. In the case of **E-G-1a**, the justification is based on demonstration that the particulate form of the NF once released into the environment experiences sufficiently fast dissolution that the role of the particle form in the toxicity of the overall NF dispersion is negligible. The particles simply do not persist for sufficient time to interact with biological interfaces in the aquatic compartment. In **E-G-1d**, whilst the NF remains to some extent intact in the particulate form, the toxicity ratio of solute to particle is sufficiently high to mean that the particles play a negligible role in the overall toxicity of the dispersion. In this instance, even though organisms are exposed to a mixture of particles and solutes in the overall exposure dispersion, it is the solutes which drive the toxicity and thus read-across to similar solutes is justified.

A simple cut-off can group NFs in **E-G-1a**, whilst in **E-G-1d**, the IATA defines the applicability range of the group, but a similarity assessment between members of the group will also be required to support the final read-across justification.

1. Identify all IATA data requirements and construct the data matrix

Whilst **E-G-1** consists of 6 sub-hypotheses, each with an associated IATA, a generic data matrix can still be generated for the overarching hypothesis. The 4 generic DNs that comprise **E-G-1** are as follows:

- DN1: Do NFs dissolve in the relevant medium / media?
- DN2: Do particles form a stable dispersion in the relevant medium / media?
- DN3: What is the density of the nanoform?
- DN4: What is the ratio of solute toxicity : particle toxicity?

The individual studies required for each tier in each DN are presented in Table 3.3.13.

*Table 3.3.13: the tiered testing strategy for each decision node (DN) required in the IATAs for the general aquatic hypothesis **E-G-1**.*

DN1: Do NFs dissolve in the relevant medium / media?	DN2: Do particles form a stable dispersion in the relevant medium / media?	DN3: What is the density of the nanoform?	DN4: what is the ratio of solute toxicity : particle toxicity?
Tier 1			
Review existing data sets			
Screening "Batch dissolution test" 24 hour single time point at a single concentration (10 mg/L); pH 5, 7 and 9 – screening for quickly dissolving NF (ISO 19057:2017, OECD GD318)	"Screening dispersion stability test" OECD TG318 media + NOM (0, 1, 10 $Ca(NO_3)_2$, pH 4, 7, 9); assessment at two time points (0 and 6 hours)	Bulk material density used as estimate of NF density	Assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests (for example, Fish, Early-life Stage Toxicity Test OECD TG210), based on comparison of generally used effect levels like EC20, EC50, LC20, LC50.
Tier 2			
Review existing data sets			
Extended "Batch dissolution test" time series in 3 pH adjusted media: pH 5, 7 and 9; 8 time points across 48 hours (OECD GD318) Extended "Continuous flow system" time series in 3 pH adjusted media: pH 5, 7 and 9; 1 mg, 12 hours, measurement as sufficient intervals to achieve constant values of dissolution rate (ISO 19057:2017, OECD GD 318)	"Extended dispersion stability test" OECD TG318 media in presence and absence of NOM (0, 1, 10 $Ca(NO_3)_2$, pH 4, 7, 9); 6 hours, assessment at hourly intervals		Extended assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests (for example: algae (OECD TG201), daphnids (OECD TG202), fish early-life stage toxicity test (OECD TG210)), based on comparison of generally used effect levels like EC20, EC50, LC20, LC50. The aim is assessing the generality of the ratio of solute : particle toxicity.
Tier 3			
Review existing data sets			
Extended "Batch dissolution test" time series in environmentally/biologically relevant media; 8 time points across 48 hours (OECD GD318) Extended "Continuous flow system" time series in environmentally/biologically relevant media; 12 hours, measurement as sufficient intervals to achieve constant values of dissolution rate (ISO 19057:2017, OECD GD 318)	"Extended dispersion stability test – heteroaggregation" OECD TG318 media in presence and absence of NOM and simulated particulate matter (0, 1, 10 $Ca(NO_3)_2$, pH 4, 7, 9); 6 hours, assessment at hourly intervals "Extended dispersion stability test" OECD TG 318 in relevant test medium/surface waters "Nanomaterial removal in wastewater" OECD WNT 3.11		Extended assessment of the ratio of solute : particle toxicity in relevant chronic tests (for example: daphnids reproduction test (OECD TG211), rainbow trout chronic toxicity on juveniles (OECD TG215)), based on comparison of generally used effect levels like EC20, EC50, LC20, LC50. The aim is assessing the generality of the ratio of solute : particle toxicity for chronic endpoints.

By combining the data requirements from both the Basic Information and the IATA a blank data matrix can be constructed (Table 3.3.14).

*Table 3.3.14: A blank simplified data matrix for the IATAs for **E-G-1***

Decision Node	Tier	Study	NF1	NF2	NF3…
Basic Information		Purity			
		Length			
		Diameter			
		Shape Aspect Ratio			
		Specific Surface Area			
		Surface chemistry			
DN1	1	Screening batch dissolution test			
		Screening continuous flow system test			
	2	Extended batch dissolution test			
		Extended continuous flow system test			
DN2	1	Screening dispersion stability test			
	2	Extended dispersion stability test			
DN3	1	Density			
DN4	1	Assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests			
	2	Extended assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests			
	3	Extended assessment of the ratio of solute : particle toxicity in relevant chronic toxicity tests			

2. Populate the data matrix with available information and identify data gaps

The available data was taken from the Basic Information and any other relevant data that had been previously generated.

*Table 3.3.15: Data matrix for the worked example for **E-G-1** populated with existing data as requested in the IATAs.*

Decision Node	Tier	Study	NF1	NF2	NF3	NF4	NF5	NF6	$AgNO_3$
Basic Information		Purity	99.9	98	99.99	99.99	46.55	90.0	>99
		Length (µm)	See diameter	See diameter	See diameter	0.1 ± 0.03	20 ± 18.1	See diameter	NA
		Diameter (nm)	2.5±1 (2-5)	20±5 (10-32)	50±35 (20-110)	20 ±3 (15 - 24)	65 ±5 (60-70)	55±8 (36-68)	NA
		Shape Aspect Ratio				5	307.7		NA
		Specific Surface Area							NA
		Surface chemistry						Ag– (Ag_3PO_4) core-(shell) configuration	NA
DN1	1	Screening batch dissolution test							Soluble
	2	Extended batch dissolution test							
		Extended continuous flow system test							
DN2	1	Screening dispersion stability test							
	2	Extended dispersion stability test							
DN3	1	Density (g/cm^3)	10.49	10.49	10.49	10.49	7.14		
DN4	1	Assessment of the ratio of solute : particle toxicity in relevant acute and/or chronic toxicity tests							Acute and/or chronic ecotoxicity data
	2	Extended assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests							
	3	Extended assessment of the ratio of solute : particle toxicity in relevant chronic toxicity tests							

In this worked example, density can already be estimated using the bulk density of the chemical (Ag for NF1 – 4 and AgO for NF5).

Reviewing existing datasets, PubChem listing of silver nitrate describes this substance as soluble, with solubility ≥ 100 gL^{-1} at 61° F. Whilst this test result comes from a different assay than the recommended screening test in DN1, it is an example of where existing data from a different assay/protocol may be used as the basis of expert judgment. With a water solubility of $AgNO_3$ ≥ 100 $g.L^{-1}$ at 61° F, it is likely to pass the 90% dissolved mass from a starting concentration of 10 mgL^{-1} which is recommended as a threshold in DN1 for quickly dissolving materials.

3. Identify potential source(s) for read-across (if required for the purpose)

There is acute and/or chronic ecotoxicity data available for silver nitrate, therefore it is a potential source for the hypothesis based on rapid dissolution of NFs (**E-G-1a**). There is no acute and/or chronic ecotoxicity data on the other candidate NFs, so if any group is formed under one of the other sub-hypotheses, it will be necessary to choose one NF on which to commission these studies and this will become the source for read-across for the other group members.

4. Examine the data requirements of each DN, including the options to move on from the DN

Examination of the exact data requirements of each DN and each tier of testing may reveal opportunities to run some studies in parallel or to avoid unnecessary testing. The user should be aware that whilst all data gaps must be addressed, there are cases where specific testing to fill a data gap is not required.

For example, if a candidate NF is found to be quickly dissolving in the tier 1 screening batch dissolution test for DN1, there is no need to assess the dispersion stability of this NF in DN2. Hypothesis E-G-1a is already fulfilled and no further assessment of the material is required, reducing the burden of testing. In this case, the user must still report why gaps in the data matrix exist. In this way the data gap is addressed but does not need specific testing as the outcome of the previous DN has waived this data requirement. This demonstrates how the sub-hypotheses in E-G-1 allows testing to be streamlined once there is sufficient evidence for a grouping outcome, rather than all DNs from the generic E-G-1 data matrix (Table 18) being required to be filled for each candidate NF.

For some DNs in the IATA, several tiers of testing are possible, and within a tier, several assays may be suggested. As a general rule, comparison between NFs should be made on results from the same assay or protocol, unless specific guidance is given on how data may be interpreted between different test systems.

As an example, two approaches are possible for the tier 2 extended dissolution testing in DN1: the batch system and the continuous flow system. Data from either may be used. It is important to note that it is not necessary to perform both tests for every candidate NF. The selection of the most appropriate method should be based on considerations of each individual NF (guidance is available in the OECD GD 318).

These decisions on testing strategy can often be made in advance to streamline assessment of the candidate NFs. Following the sequential order of the DNs and the tiered testing strategy presented in the IATAs will allow for NFs to be tested only to the extent required to come to a grouping decision, rather than complete assessment for all gaps in the data matrix for all candidate NFs.

5. **Perform study(ies) within tier 1 of DN1 in the IATA**
6. **Decide whether DN1 is sufficiently addressed by tier 1 or whether higher tier studies are needed, then perform higher tier studies if required.**
7. **Repeat stages 6 and 7 for each relevant tier for each DN until all DNs have been addressed**

This worked example takes the user through all the studies in each DN showing how the results of the studies in the IATA have been used to decide how to progress both through the tiers in a DN and through the DNs to complete the data matrix. Therefore stages 5, 6 and 7 are all addressed concurrently.

Full details of the experimental work and decision-making process can be seen in (Song et al., in preparation). A summary of each DN is given here.

DN1: Do NFs dissolve in the relevant medium/media?

This DN identifies NFs which can be defined as quickly dissolving **(E-G-1a**), partially dissolving (**E-G-1d – f**) and very slowly dissolving (**E-G-1b – c**). Only those NFs identified as partially or very slowly dissolving need to progress to DN 2 in the IATA.

After reviewing existing data all 6 candidate nanoforms require testing to fill the gap in the data matrix.

Tier 1: Screening batch dissolution test.

This screening test is described in the tiered testing strategy for **E-G-1**. The screening batch dissolution test addresses a single time point after 24 hours under three conditions of pH (4, 7 and 9), with a starting concentration of the material of 10 mg/L.

The threshold for "quickly dissolving" NFs is for ≥90% of the NF mass to be dissolved under these three conditions after 24 hours.

Table 3.3.16: Percentage of NF dissolved (starting with 10mg/L) after 24 hours, measured for the six candidate nanoforms for DN1, tier 1 screening dissolution testing in IATA E-G-1 (Zhang, 2011).

NF	% dissolved at 24 hours		
	pH4	pH7	pH9
NF1	99	95	91
NF2	30	20	15
NF3	12	10	8
NF4	8	5	6
NF5	5	5	<LOD
NF6	<LOD	<LOD	<LOD

From the screening batch dissolution test (Table 3.3.16), **NF1** passes the threshold for quickly dissolving NFs. Further assessments for NF1 should consider exposure to solutes only. No progression to higher tier tests in DN1 is required. Consideration can be made whether the solutes remain in their form, without chemical transformation into a new NF. However, as the solutes from this NF would be silver ions (purity was 99.9%, no significant impurities expected), the fate of the dissolution products from NF1 would be expected to follow a similar chemistry as the solutes of $AgNO_3$ when in a similar media. Therefore, NF1

may be grouped according to hypothesis **E-G-1a** and this successful grouping outcome may be used as part of a read-across justification to the existing source data for $AgNO_3$.

None of the remaining candidate NFs pass the quickly dissolving threshold for this screening batch dissolution test. This triggers progression to the higher tier 2 extended dissolution testing for NFs2-6.

Tier 2: Extended dissolution testing

Two approaches are possible for the extended dissolution test: the batch system and the continuous flow system. In this instance, information from both approaches can be considered as equivalent data from the same assay. The conditions and interpretation of the data is the same for both approaches, it is simply the test set-up which differs as the two approaches are suitable for different ranges of material solubility. The selection of the most appropriate method should be based on considerations of each individual NF. Guidance on how to select either the batch or continuous flow system is detailed in the OECD GD 318. Whilst exact cut-offs for applying a continuous flow test versus a batch test cannot be given, it is suggested that the continuous flow test is most appropriate if the solubility of the nanomaterial is between 0.1 and 10 mg/L, whilst a batch test is perhaps more appropriate for NFs with a solubility <1 mg/L (OECD GD 318).

In this example, **NF2** and **NF3** would be best suited to the continuous flow system, whilst **NF4**, **NF5** and **NF6** would be better suited to the extended batch test.

*Table 3.3.17: Dissolution rates (ng/cm²/h) measured on the six candidate nanoforms for DN1 according to tier 2 extended dissolution testing in IATA **E-G-1**.*

NF	Test method	Dissolution rates (ng/cm²/h)		
		pH4	pH7	pH9
NF1	Not tested	Not tested	Not tested	Not tested
NF2	Continuous flow	350	300	250
NF3	Continuous flow	175	160	150
NF4	Batch	340	310	265
NF5	Batch	150	125	100
NF6	Batch	<LOD	<LOD	<LOD

To demonstrate the extended dissolution testing, Table 3.3.17 summarizes the dissolution rates calculated as a function of the total surface area of the particles in the tests.

Tier 3: Extended dissolution testing in specific media

As the purpose of grouping is to demonstrate similarity between candidate NFs as part of a read-across justification to the source data for higher tier ecotoxicity tests for the soluble source material $AgNO_3$, data from tier 2 is considered sufficient for this purpose. Escalation to tier 3 could be considered if it is beneficial to demonstrate similarity in dissolution behavior between NFs in a specific ecotoxicity media, or a natural surface water for which a risk assessment is targeted.

Conclusion from DN1

NF1 is considered quickly dissolving on the basis of the tier 1 screening batch test for dissolution. For this NF, hypothesis **E-G-1a** is accepted and no further testing is required.

NFs2 - 5 undergo incomplete dissolution in the extended test after 48 hours. Hypotheses **E-G-1d, e and f** are now considered for NF2 - 5. These NFs all fall within the groups considered to be partially dissolving according to E-G-1. They neither conform to the threshold for quickly dissolving NFs (from the tier 1 screening test), nor the very slowly dissolving threshold (potential dissolved fraction is below the limit of detection). These two thresholds define the applicability range for particles defined as partially dissolving. This triggers hypotheses **E-G-1d, e and f** to be considered. Similarity assessment of the dissolution rates from this DN can be used in conjunction with outcomes from DN3 (dispersion stability), DN4 (density) and DN5 (ratio of solute to particle toxicity) to arrive at a final grouping decision for these 4 candidate NFs. Similarity assessment between the dissolution rates of NFs2-5 should be used as part of the justification for read-across if these materials are grouped successfully in later stages of the IATA. More details on performing a similarity assessment on individual properties can be found in section 3.3.5. Dissolution products of NF6 could still not be detected above the limits of detection in the extended batch test in tier 2, indicating very slow dissolution. Hypotheses **E-G-1b** and **E-C-1c** are triggered for NF6 as it is now considered a slowly dissolving NF on the basis of DN1.

DN2: Do particles form a stable dispersion in the relevant medium / media?

Only candidate NF1 was identified as quickly dissolving in DN1. This NF requires no further assessment in **E-G-1**. The remaining candidate NFs require assessment of the dispersion stability of these materials, following the tiered testing strategy in Table 3.3.14.

Complete guidance on performing the screening dispersion stability test can be found in OECD TG 318. The principle is that the screening dispersion stability test assesses stability in a matrix of test media representing different combinations of pH and $Ca(NO_3)_2$ concentrations in the presence of natural organic matter at two time points (0 and 6 hours). This can screen for NFs that form a stable dispersion and those that form an unstable dispersion, neither of which require escalation to tier 2 extended testing.

For particles which are found to be partially stable (note: in the terminology of OECD TG 318 this is a dispersion of intermediate stability), extended testing can allow for more detailed assessment of how the NFs are behaving in the test, for example distinguishing between particles which form stable aggregate/agglomerate (red line Figure 3.3.10), those which represent a NF that continuously agglomerates and settles out of dispersion (yellow line) and a NF that quickly agglomerates and settles or contains two different fractions, a stable and unstable fraction (green line). The results from this extended test are thus useful for similarity assessment of this property between NFs.

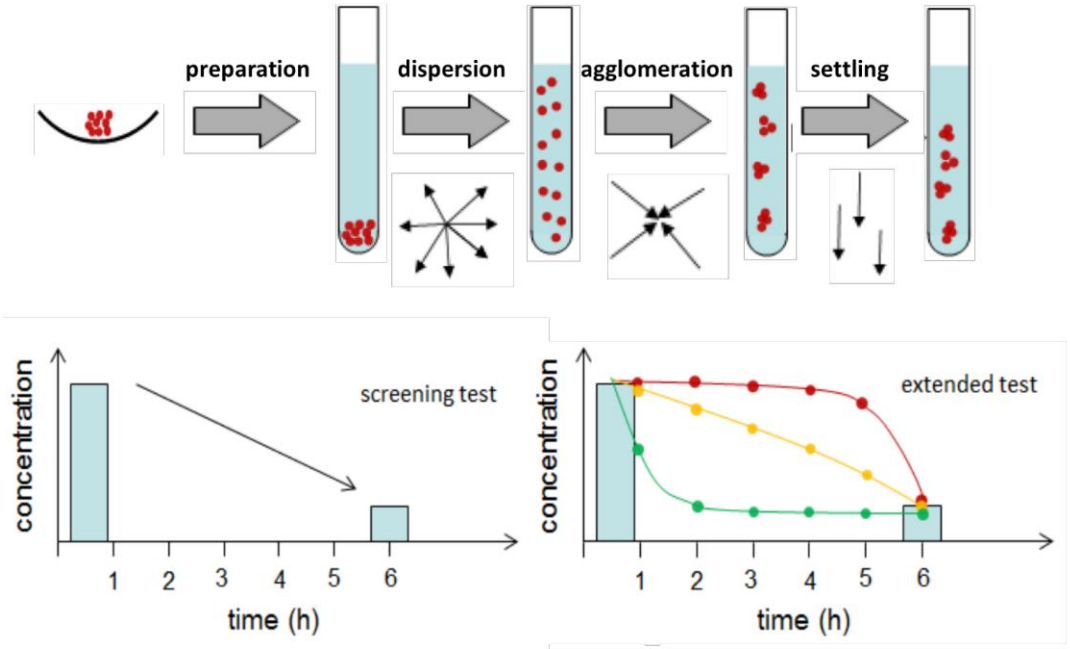


Figure 3.3.10: Figure and caption text reproduced from OECD TG 318 "Dispersion stability of nanomaterials in simulated environmental media". Principle of the testing scheme (upper part), and possible outcome of the screening and the extended testing (lower part; see text and figure 2). The red line represents a nanomaterial that has a small density difference to water, it agglomerates, but almost does not settle. In step 6, a centrifugation is performed which takes the density of the nanomaterial into account and performs a particle size cut-off. Here, the lightweight agglomerates are removed into the deposited fraction because they are larger than the cut- off value. The yellow line represents a nanomaterial that continuously agglomerates and settles out. The green line represents either a nanomaterial that quickly agglomerates and settles (high density) or a heterogeneous nanomaterial that contains two different fractions. For the first, the number concentration in the top of the vial is reduced so that the further agglomeration is slowed down to a point where it becomes virtually stable (not enough collisions in the timeframe). For the latter, one fraction is unstable and settled out within 2 h, another fraction is highly stable and does not agglomerate.

Tier 1: Screening dispersion test.

NF1 does not require testing by DN2 as it has been identified as quickly dissolving.

NF2 is identified as forming a stable dispersion (Table 3.3.18). Across the matrix of treatments, varying $Ca(NO_3)_2$ and pH, ≥ 90% of the material remained in dispersion after 6 hours. No further testing in DN2 is required for NF2.

NF6 is identified as forming an unstable dispersion. Across the matrix of treatments, varying $Ca(NO_3)_2$ and pH, ≤ 10% of the material remained in dispersion after 6 hours. No further testing in DN2 is required for NF2.

NFs 3, 4 and 5 require escalation to tier 2, extended dispersion stability testing to resolve information on the trend of settling and understand the underlying sedimentation process as part of a similarity assessment between NFs grouped as partially stable.

*Table 3.3.18 Data measured on the six candidate nanoforms for DN2, tier 1 screening dispersion stability testing in IATA **E-G-1**. Values reported are the percentage of initial mass remaining in dispersion at the end of the 6 hour screening test.*

NF 1	$Ca(NO_3)_2$		
pH	0	1	10
4	Not tested		
7			
9			

NF 2	$Ca(NO_3)_2$		
pH	0	1	10
4	99	98	95
7	95	94	90
9	96	95	92

NF 3	$Ca(NO_3)_2$		
pH	0	1	10
4	95	95	91
7	93	92	91
9	95	94	91

NF 4	$Ca(NO_3)_2$		
pH	0	1	10
4	70	68	55
7	50	41	8
9	56	45	35

NF 5	$Ca(NO_3)_2$		
pH	0	1	10
4	30	25	15
7	20	12	6
9	22	15	11

NF 6	$Ca(NO_3)_2$		
pH	0	1	10
4	4	2	1
7	1	0.2	0.1
9	2	1	0.5

Tier 2: Extended dispersion test.

Figure 3.3.11 illustrates how the extended dispersion stability test can give additional insights into the underlying sedimentation process. The screening at 6 hours would indicate all three NFs (NF3, 4 and 5) have a similarly low stability (all < 50% material remaining in dispersion after 6 hours; Table 3.3.18), but Figure 3.3.11 shows that the underlying process of sedimentation is different for NF5 as compared to NF3-4. It should be noted that the graph is for illustrative purposes and only demonstrates the relationship between dispersion stability and time for a single test condition (i.e. pH 7 and 10 mM $Ca(NO_3)_2$).

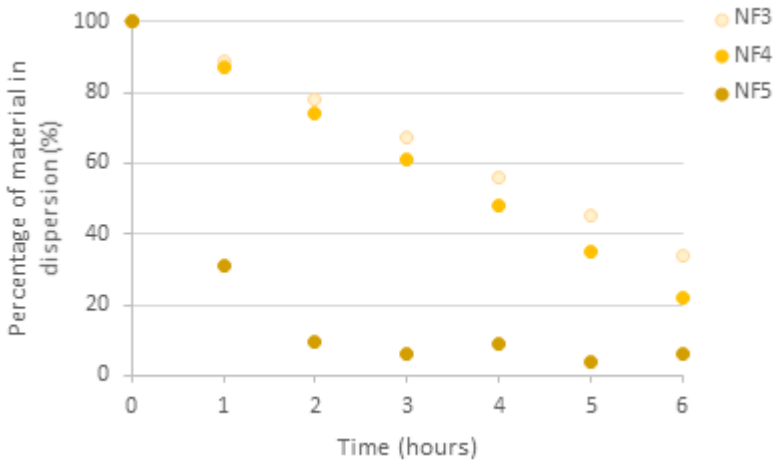


Figure 3.3.11: An example of the data generated for the extended dispersion stability test for the three partially stable NFs identified in the tier 1 screening dispersion stability assay. Data represents the dispersion stability under a single condition, pH 7, 10 mM $Ca(NO_3)_2$, as an illustrative example of how the extended test can identify differences in the underlying sedimentation process for the three NFs.

To demonstrate similarity between NF3 and NF4, it may be sufficient to perform assessment on a scalar descriptor for dispersion stability, such as the half time for 50% of the material to have settled out of dispersion, or the percentage of material remaining in dispersion at a given time point (see section 3.3.5). This is because both materials seem to experience a similar pattern of steady settling of material over time at a continuous rate.

To demonstrate the dissimilar behavior between NF5 as compared to NF3 and NF4, comparison of the curves is required. More details on similarity assessment of different types of data can be found in Jeliazkova *et al.* (submitted NanoImpact Similarity Special Issue). In this example, NF3 and NF4 appear to be characterised by a steady decrease in mass of original particles remaining in dispersion, whilst NF5 is characteristic of a heterogeneous material in which two fractions exist. The Basic Information indicates that NF5 is a polydisperse population of nanowires, with a high standard deviation in lengths of 20 ± 18.1 µm. As such, the curve illustrated in Figure 20 might be indicative of larger, less stable nanowires settling rapidly out of dispersion in the first few hours, with a second fraction of ~10% the original starting mass, remaining stable in dispersion. This second fraction could be closer in characteristics to the elongated nano-rods of NF4 for example.

Tier 3: Extended dispersion stability testing in specific media

As the purpose of grouping is to demonstrate similarity between candidate NFs as part of a read-across justification between members of a group defined as partially stable, data from tier 2 is considered sufficient for this purpose.

Escalation to tier 3 could be considered if it is beneficial to demonstrate similarity in dispersion stability behavior between NFs in a specific ecotoxicity media, if the purpose of read-across is for a specific ecotoxicity endpoint, or a natural surface water for which a risk assessment is targeted.

Conclusion from DN2

NF2 is identified as forming a stable dispersion. Hypotheses **E-G-1d, e and f** are triggered for this NF, with the material considered partially dissolving and forming a stable dispersion.

NF3, 4 and 5 are identified to be partially stable dispersions. Hypotheses **E-G-1d, e and f** are triggered for these NFs, with the material considered partially dissolving and forming a partially stable dispersion. Whilst all three NFs fall within the applicability range of these hypotheses (being neither stable, nor unstable dispersions), further assessment of similarity on the tier 2 data from DN2 should be taken into consideration when justifying a read-across decision between these particles.

NF6 is identified as forming an unstable dispersion. It is no longer considered under **E-G-1**. Hypotheses **E-WS-1a to e** should be considered concerning sediment exposures to NF.

DN3: Is the density of the NF <1 g/cm^3?

The density of each candidate NF is estimated from the bulk density of the material. For example, NF5 is an Ag_2O nanowire. The density of silver oxide (7.14 g/cm^3) is used as an estimate of the density of the NF.

All NFs are considered to be denser than water, and so sedimentation and exposure of sediments over extended time scales should also be considered as a relevant endpoint. No testing is required in this case for any of the NFs.

Conclusion from DN3

Proceed to **DN4**

DN4: What is the ratio of solute toxicity : particle toxicity?

Two tests are needed to determine the ratio: one with NFs in which release of dissolution products is monitored over time, and another with dissolution products only. Within the tiers differences between species may occur: expert judgement may be needed to decide on the most relevant species for the grouping purpose.

NF1 and NF6 do not require testing in this DN. NF1 is already considered quickly dissolving (tier 1 DN1), identifying the solutes as responsible for driving toxicity. NF6 is already considered very slowly dissolving (tier 2, DN2) and so the solutes are not considered to contribute to the overall exposure.

NFs 2, 3, 4 and 5 require assessment of the ratio of solute toxicity to particle toxicity to demonstrate whether read-across to the solute is still justified (**E-G-1d**) or if read-across to a similar NF is more appropriate (**E-G-1e**).

Tier 1: Assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests

The relevant tests have been described in general terms in Table 10. The contribution to suspension toxicity of the solutes compared to the contribution of the particles to the overall suspension toxicity for any NF depends on the following properties:

1. Dissolution rate: It is to be noted that steady state conditions are not considered. This is because at a given steady state concentration, the assessment of the ratio of toxicity of solutes compared to particles is not only relatively straightforward (as this can be done by means of a direct comparison of ECx or LCx values, assuming similarity of the dose-response curves of solutes and particles), but also there is no possibility to predict the solubility of NFs for a specific test medium.
2. Initial particle concentration (ion concentration is assumed to be 0 at the start of the exposure).
3. Particle size.
4. Toxicity of the ions and toxicity of the NFs on the basis of the available ECx or LCx values.
5. The shape of the dose-response curve of particles and ions, as can be quantified by means of the Hill coefficient (Hill, 1910). Significantly different dose-response curves imply that at different concentrations of solutes and NFs, the contribution of either ions or particles to overall suspension toxicity might shift. When no information on the shape of the dose-response curve is available, it can be assumed that the shapes are similar for ions and NFs.
6. Test duration. The dose-response curves of ions and NFs depend on the test duration. For convenience it can be assumed that the Hill-coefficient is independent of exposure duration.

The aim is to quantify which ratio(s) of ion toxicity : particle toxicity, in connection with dissolution rate, can be used as:

- a cut-off above which ion toxicity is appropriate as a proxy to quantify the toxicity of suspensions of particles, or
- a cut-off for deciding below which ions do not contribute significantly to toxicity.

The Hill equation may be used in its simplest form to assess the toxicity of either ions or particles at a specific time point:

$$E = \frac{100}{1+\left(\frac{EC_{50}}{[A]}\right)^{n_H}}$$

In this equation, n_H is the Hill coefficient. When no specific information is available, a value of 6 may be taken as being a typical value for Ag NFs and Ag ions. E is the observed effect, A is the concentration of either NFs or ions. Commonly, dose response curves will be experimentally determined separately for ions and for suspensions of NFs. In the latter case, the toxicity of the NFs in the test medium can be deduced from the measured suspension assuming mixture toxicity (so-called response addition):

$$E_{total} = 1 - \left[\left(1 - E_{Ag+}\right)\left(1 - E_{AgNF}\right)\right]$$

In this equation, E_{total} is the measured suspension toxicity, E_{Ag+} is the measured toxicity of the Ag-ions present in suspension, and E_{AgNF} is the toxicity of the Ag NF as determined with any of the tiered tests mentioned in Table 10. Although no formal cut-off values have been agreed upon yet, tentative cut-off values of 9:1 and 1:9 in terms of the effect caused by ions and a NF can initially be selected as cut-off

values for **E-G-1d** and **E-G-1e** (or: at least 90 % of the effect due to ions in case of DN **E-G-1d** and at least 90 % of the effect due to the NF in case of DN **E-G-1e**). The results of the toxicity measurements are shown in Table 3.3.19.

Table 3.3.19. Toxicity ratios *measured on four of the six candidate nanoforms for DN4.*

NF	Test method	Measured ratio of solute : particle toxicity in relevant toxicity test
NF1	Not tested	Not tested
NF2	Daphnids reproduction test (OECD TG211)	95 : 5
NF3	Daphnids reproduction test (OECD TG211)	90 : 10
NF4	Daphnids reproduction test (OECD TG211)	4 : 96
NF5	Daphnids reproduction test (OECD TG211)	1 : 99
NF6	Not tested	Not tested

Conclusions from DN4

NF2 and 3 are identified as NFs for which the toxicity is dominated by the toxicity of the ions. Ion-specific toxicity data are generally sufficient for the requirements of each of the tiers indicated in Table 3.3.19 (although some caution may be needed, see next paragraphs).

NF4 and 5 are identified as NFs for which suspension toxicity is dominated by toxicity of the NFs. NF-specific toxicity data are required according to the requirements of each of the tiers indicated in Table 3.3.19.

It is to be noted that it is common practice in chronic toxicity testing to regularly replace the test medium across the test duration. The behavior and state of exposure of NF suspensions in a short term test can therefore be considered representative of the behavior and state of exposure of these NFs in longer term or chronic studies.

If only because this is out of the scope of the tiered testing proposed here, we do not make specific recommendations about the predictivity of lower tier tests to higher tier toxicity tests. However, where some comparative toxicity data (e.g. Daphnids vs. fish) for NFs in a group are available, arguments over consistent exposure across tests could be considered and may suffice as (part of) a read-across justification for other NFs.

The populated data matrix for the six candidate NFs following the IATA for **E-G-1** is presented in Table 3.3.20.

Table 3.3.20: Example of a simplified populated data matrix for the six candidate NFs following the IATA for E-G-1

Decision Node	Tier	Study	NF1	NF2	NF3	NF4	NF5	NF6	$AgNO_3$
Basic Information		Purity	99.9	98	99.99	99.99	46.55	90.0	>99
		Length (µm)	See diameter	See diameter	See diameter	0.1 ± 0.03	20 ± 18.1	See diameter	NA
		Diameter (nm)	2.5±1 (2-5)	20±5 (10-32)	50±35 (20-110)	20 ±3 (15 - 24)	65 ±5 (60-70)	55±8 (36-68)	NA
		Shape Aspect Ratio				5	307.7		NA
		Specific Surface Area							NA
		Surface chemistry	Not tested	Not tested	Not tested	Not tested	Not tested	Ag– (Ag_3PO_4) core-(shell) configuration	NA
DN1	1	Screening batch dissolution test	≥90%	<30%	<12%	<8%	<5%	<LOD	Soluble
	2	Extended batch dissolution test	Not required	Not required	Not required	Partial	Partial	<LOD	Not required
		Extended continuous flow system test	Not required	Partial	Partial	Not required	Not required	Not required	Not required
DN2	1	Screening dispersion stability test	Not required	>95%	<65%	<70%	<30%	<4%	Not required
	2	Extended dispersion stability test	Not required	Not required	Partial	Partial	Partial	Not required	Not required
DN3	1	Density (g/cm^3)	10.49	10.49	10.49	10.49	7.14	6.37	4.35
DN4	1	Assessment of the ratio of solute : particle toxicity in relevant acute toxicity tests	Not tested	Higher than 90 %	Higher than 90 %	Lower than 10 %	Lower than 10 %	Not required	Acute ecotoxicity data

8. Use the results from all the DN to identify a grouping conclusion

The initial hypothesis was that candidate silver NFs could be grouped as part of a read-across justification to existing data for the soluble ion of silver, using $AgNO_3$ as a source material for the ecotoxicity of this solute. This grouping hypothesis was to be on the foundation of demonstrating similar behaviors in waters that lead to the solute of the material driving the toxicity of the NF dispersion and so forming part of the justification that read-across to the existing data for soluble $AgNO_3$ would be a conservative estimate of the aquatic toxicity of these candidate NFs.

For three of the candidate NFs, NF1, 2 and 3, grouping using the IATA for **E-G-1** provides evidence that supports read-across to the existing data for the dissolved silver ion. It should be noted that whilst the identified source material ($AgNO_3$) is the same for these three NFs, the justification behind this conclusion differs.

NF1 was found to be quickly dissolving and so grouped under hypothesis **E-G-1a**. This hypothesis concludes that the rapid dissolution of the NF in the environment would lead to a negligible exposure of particles to biota. Therefore, it is only the solute which is considered to interact with organisms and thus read-across to a soluble source material is justified.

NF2 and 3 were found to be partially dissolving but with a similar stability to the particle populations in aquatic media. Similarity assessment of the tier 2 data for both dissolution and dispersion stability demonstrates the behavior of these two NFs is similar. The ratio of toxicity of solutes to particles for both NFs was also high, indicating that the solutes are primarily responsible for the toxicity of the overall NF suspension (**E-G-1d**). This can be used as evidence to support a read-across justification to the existing data for the soluble form of silver on the basis that the solutes were driving effects for these NFs, limited

only by the dissolution rates of the materials. The contribution of the soluble form of Ag in suspensions of NF2 and 3 is therefore significantly higher than the contribution to suspension toxicity of the two NFs, and so would be a conservative estimate of the toxicity for these materials.

NF4 and 5 were both found to be partially dissolving and partially stable particles. Whilst their rates of dissolution were found to be similar, tier 2 extended dispersion stability testing identified differences in the underlying process of sedimentation between the two. Both NFs were also found to have a low solute to particle toxicity ratio and so were grouped under hypothesis **E-G-1f**. This means that it is the particles driving the toxicity of the NF suspensions, not the solutes. The significance of the difference in dispersion stability would therefore have to be carefully considered as part of a similarity assessment on a property-by-property basis, as this could influence the toxicity of these two materials, resulting in a difference in the exposure of the NFs to organisms. However, provided justification is given, similarity assessment could conclude that these two NFs are similar enough to read-across data from one to fill gaps in ecotoxicity testing for the other. In this instance, testing could be undertaken for one material and then this data used for read-across to the other.

NF6 was found to be very slowly dissolving and particles quickly removed from the aquatic environment (due to the low dispersion stability). This would indicate a low toxicity ratio of solutes to particles due to the low dissolution of silver from this NF. For aquatic testing, alternative similar source materials would need to be identified, which would require similar chemistry to this core-shell structure of Ag-(Ag_3PO_4), very slow dissolution and low dispersion stability. Alternatively, specific testing is recommended for this NF. It is also recommended that sediment exposure is considered for this material as it is identified as being unstable in the aquatic environment. This may be explored using the IATA for **E-WS-1**.

Overall Conclusions:

The summarised conclusions to the grouping are shown in Table 3.3.21.

*Table 3.3.21: summary of the grouping outcomes for each of the six candidate NFs in **E-G-1** on the basis of the results generated across the decision nodes in the IATA.*

NF	DN1	DN2	DN3	DN4	Grouping outcome	Identified source material for read-across
NF1	Quickly dissolving	Not required	Not required	Not required	E-G-1a	$AgNO_3$
NF2	Partially dissolving	Stable dispersion	No, density is >1 g/cm^3	Ratio of solute : particle toxicity is high	E-G-1d	$AgNO_3$ or other candidate NF in E-G-1d based on similarity assessment
NF3	Partially dissolving	Partially stable dispersion	No, density is >1 g/cm^3	Ratio of solute : particle toxicity is high	E-G-1d	$AgNO_3$ or other candidate NF in E-G-1d based on similarity assessment
NF4	Partially dissolving	Partially stable dispersion	No, density is >1 g/cm^3	Ratio of solute : particle toxicity is low	E-G-1f	Similarity assessment on individual properties between NF4 and NF5 may justify assessment of one NF and read-across to the other.
NF5	Partially dissolving	Partially stable dispersion	No, density is >1 g/cm^3	Ratio of solute : particle toxicity is low	E-G-1f	Similarity assessment on individual properties between NF4 and NF5 may justify assessment of one NF and read-across to the other.
NF6	Very slowly dissolving	Unstable dispersion	No, density is >1 g/cm^3	Not required	E-G-1c (E-WS-1 IATA is also relevant)	Alternative similar source NF must be found or testing of this NF is recommended

3.3.3.4 Do you want to know more?

These references include those used in the previous section but also include other sources of information that expands on that given in this Guidance Document.

eNanomapper (http://www.enanomapper.net/)

> *eNanoMapper provides a computational infrastructure for toxicological data management of engineered nanomaterials (ENMs) based on open standards, ontologies and an interoperable design to enable a more effective, integrated approach to European research in nanotechnology.*

Hill, A.V. (1910) The possible effects of the aggregation of the molecules of hemoglobin on its dissociation curves. J. Physiol. (Lond), 40, 4-7.

> *Describes the development of the Hill co-efficient, used in the assessment of toxicity of ions or particles.*

Jeliazkova *et al.* (submitted NanoImpact Similarity Special Issue)

Murphy F., Schinwald A., Poland C.,Donaldson K. (2012). The mechanism of pleural inflammation by long carbon nanotubes: interaction of long fibres with macrophages stimulates them to amplify pro-inflammatory responses in mesothelial cells. Part Fibre Toxicon., 9: 8, doi: 10.1186/1743-8977-9-8

> *Presents experimental data to support the hypothesis that long fibres elicit an inflammatory response in the pleural cavity via frustrated phagocytosis in pleural macrophages.*

Murphy F., Dekkers S., Braakhuis H., Ma-Hock L, Johnston H., Janer G., di Cristo L., Sabella S., Jacobsen N., Oomen A., Haase A., Fernandes T., Stone V. (2021). An integrated approach to testing and assessment of high aspect ratio nanomaterials and its application for grouping based on a common mesothelioma hazard. NanoImpact, 22, 100314, doi 10.1016/j.impact.2021.100314.

> *This paper explains how the hypotheses and IATA were developed and gives in depth detail around the experimental work to justify the hypothesis.*

OECD TG 318. Guidance document for the testing of dissolution and dispersion stability of nanomaterials and the use of the data for further environmental testing and assessment strategies.

> *Standardised method used to measure dissolution rate and dispersion stability of nanoforms. This is the recommended study to fulfil these endpoints for REACH registrations.*

Palomäki J., Välimäki E., Sund J., Vippola M., Clausen P., Jensen K., Savolainen K., Matikainen S., Alenius H. (2011). Long, needle-like carbon nanotubes and asbestos activate the NLRP3 inflammasome through a similar mechanism. ACS Nano, Sep 27;5(9):6861-70. doi: 10.1021/nn200595c

Investigates whether different carbon nanomaterials induce a pro-inflammatory response in human primary macrophages

Roebben G., Rasmussen K., Kestens V., Linsinger T., Rauscher H., Emons H., Stamm H. (2013). Reference materials and representative test materials: the nanotechnology case. J Nanopart Res, 15, 1455, doi 10.1007/s11051-013-1455-2.

Provides an overview of the existing types of reference materials and introduces a new class of test materials for which the term 'representative test material' is proposed. Illustrates this system with examples from the field of nanomaterials, including reference materials and representative test materials developed at the European Commission's Joint Research Centre

Song et al. (in preparation for submission) . Similarity assessment of metallic nanoparticles within a risk assessment framework: a case study on metallic nanoparticles. NanoImpact Similarity Special Issue, 202

A case study performed aimed at assessing the similarity of a set of spherical metallic NFs that different with regard to chemical composition and particle size. The endpoints of assessment were root elongation and biomass increase of lettuce seedlings.

World Health Organisation (1996). Determination of Airborne Fibre Number Concentrations. A Recommended Method, by Phase Contrast Optical Microscopy (Membrane Filter Method), World Health Organization, Geneva

Defines the particle characteristics of respirable fibres that present a significant health risk

Zhang W., Yao Y., Sullivan N., and Chen Y. (2011). Modeling the Primary Size Effects of Citrate-Coated Silver Nanoparticles on Their Ion Release Kinetics. Environ. Sci. Technol., 45, 10, 4422–4428. doi.org/10.1021/es104205a

Provides fundamental insight into the ion release kinetics of AgNPs in aqueous environments, allowing improved understanding and predicting the nanotoxicity of AgNPs.

3.3.4 Use of Data Quality Assessment within the GRACIOUS Framework

A scientifically justified grouping decision must be made using physicochemical and (eco)toxicological data in which there is sufficient confidence in their quality. This applies both to existing data extracted from external sources and for new data. The GRACIOUS Framework recommends the use of a "traffic light" approach to assessing the overall quality of such data.

This approach automates the data quality assessment process starting from the available (meta)data. It requires minimal expert judgment and has been implemented in the eNanoMapper database to enable real-time analysis of each dataset that is included in it.

3.3.4.1 Instructions to assess data quality

The data quality assessment approach is based on four established criteria, namely:

> **1. Data completeness**: which refers to the degree to which all required (meta)data in a data set is available;
>
> **2. Data reliability**: which measures if a study was conducted in a reliable manner;
>
> **3. Data relevance**: which measures if a study was conducted using agreed (standard) protocols/procedures;
>
> **4. Data adequacy**: defining the usefulness of the data for risk assessment purposes.

Data completeness

Data completeness is evaluated with respect to an (eco)toxicological endpoint of interest, ensuring that both a proper physicochemical characterization and sufficient information related to the testing procedure and test conditions have been provided. A Completeness Score (CS) is computed as the number of items (parameters) reported in a data entry template divided by the number of items (parameters) required by the template (Comandella *et al.*, 2020). We have applied this approach to NANoREG and GRACIOUS data entry templates (Gottardo *et al.*, 2019), which are being used for data entry in many EU funded projects and are currently being standardized.

Data reliability

Following the works of Card & Magnusson (2010) and Fernandez-Cruz *et al.* (2018), data reliability assessment is performed using the ToxRTool. ToxRTool is an excel spreadsheet, which assigns data to the following categories:

> **1. Category 1** - Reliable without restriction: Studies or data from the literature or reports which were carried out or generated according to generally valid and/or internationally accepted testing guidelines or in which the test parameters documented are based on a specific (national) testing guideline or in which all parameters described are closely related/comparable to a guideline method.
>
> **2. Category 2** - Reliable with restrictions: Studies or data from the literature, reports in which the test parameters documented do not totally comply with the specific testing guideline, but are sufficient to accept the data or in which investigations are described which cannot be subsumed under a testing guideline, but which are nevertheless well documented and scientifically acceptable.

3. Category 3 - Not reliable: Studies or data from the literature/reports in which there were interferences between the measuring system and the test substance or in which organisms/test systems were used which are not relevant in relation to the exposure (e.g., non-physiological pathways of application) or which were carried out or generated according to a method which is not acceptable, the documentation of which is not sufficient for assessment and which is not convincing for an expert judgment.

It does not assign to category 4 ("Not assignable"), as this should be made by direct consideration of the user.

Data relevance

Relevance covers the extent to which data and test are appropriate for hazard characterization. Specifically, this step ensures that the study is conducted using protocols and procedures that are relevant to identify the hazards related to the endpoint.

Four tentative categories have been defined for data relevance:

1. Category 1: data derived by means of internationally recognized standard guidelines, such as the OECD TGs (guidelines must be nanospecific or applicable to nanoforms).

2. Category 2: derived using nanospecific validated protocols, and protocols that are candidates to become OECD TGs or OECD TGs with modifications. Validation of protocols should have been performed using internationally accepted principles and procedures (Magnusson and Örnemark, 2014) (eds.)

3. Category 3: data for which the protocol is not included in categories 1 and 2 (including nanospecific protocols which are not yet validated).

4. Category 4: data for which the adopted protocol is not reported.

Data adequacy

Adequacy defines the usefulness of data for the purposes of the analysis. Three main types of studies were selected, namely *in vivo, in vitro* and *in silico*. Usually, higher weight is associated to the most reliable test for risk assessment purposes (i.e., *in vivo)*, while lower weights were associated to *in vitro* and *in silico* studies, unless where specific *in vitro* methods are actively encouraged or required by regulators.

Assessing data quality

Scores are computed for each of the above criteria and then aggregated into an overall data quality score (this work will be finalised in the Gov4Nano project (https://www.gov4nano.eu/). The calculated score is then assigned to a specific light in the "traffic light" system based on pre-defined thresholds. The whole process has been automated, thus highlighting data quality and completeness directly on the database user interface and/or in the data reporting templates when uploading/downloading data from the database. The data quality assessment approach has been implemented as functionality of the eNanoMapper database (Outside eNanoMapper, the quality assessment can be applied if the data allow assessment of the criteria. For example, Completeness should be defined and be comparable to a "reference" and for the other criteria the information that allows their assessment must be available, so in principle it could be implemented for other databases

3.3.4.2 Example: Assessing data quality for a quantitative Weight of Evidence approach for hazard classification of nanomaterials according to the EU CLP Regulation

A case study of the assessment of data quality based on the criteria presented in the previous section is included in Basei *et al.* (submitted). In that example, quantitative assessment of data quality is used in a weight-of-evidence approach for hazard classification of nanomaterials according to the CLP Regulation.

In brief, the data quality criteria completeness, reliability, relevance and adequacy are evaluated for data related to physicochemical properties plus the (eco)toxicity endpoint "aquatic toxicity" extracted from the eNanoMapper database. The results of the evaluation of the different quality criteria are translated (for each quality criterion) into numerical values. These values are then used as weights of each study for further analysis.

In this example, extraction and curation of the eNanoMapper data as described in (Basei et al) reduces the information to 47 entries of data on aquatic toxicity, generated by the NANoREG and MARINA projects, related to 13 NMs. These entries are related mostly to acute Aquatic Toxicity (45 entries), while two studies address chronic Aquatic Toxicity.

Completeness Scores (CS) are computed for each study based on a checklist of properties and conditions acquired from eNanoMapper templates, which are refined versions of the NANoREG and GRACIOUS templates. For each study the CS is computed for each relevant template related to the physicochemical characterization as the number of items reported, divided by the number of items required by the template. The following 11 physicochemical properties were considered: crystallinity, composition, particle size, surface chemistry, particle shape, specific surface area, surface charge, surface hydrophobicity, dustiness, water solubility and density (Comandella, 2020). In addition, for (eco)toxicological datasets (here aquatic toxicity) the evaluation of data completeness covers the information related to the testing procedure (e.g. reference to the Standard Operating Procedure, the tested endpoint, the assay name, etc.) and test conditions (e.g. the adopted dispersion protocol and medium, the concentration, details on the cell lines and culture conditions, etc.).

The CSs of the 11 physicochemical parameters are then averaged, obtaining a score for the study related to the physicochemical characterization of the NM, $CS_{physchem}$. The CS associated to the template of the (eco)toxicological endpoint, CS_{ecotox}, is computed analogously, and finally $CS_{physchem}$ and CS_{ecotox} are averaged, thus obtaining an overall CS for a particular study quantifying the completeness of the information related both to the physicochemical characterization of the NM and the characterization of (eco)toxicological study associated to the endpoint of interest:

The overall CS of a specific study is then used as weight of that study for further analysis.

For the quality criteria reliability, relevance and adequacy numerical values based on the respective categories are defined, to be used as weights for further analysis.

For assessment of the data **reliability** the numerical values to be used as weights are defined as:

- Category 1: 1
- Category 2: 0.5
- Category 3 or 4: 0

For assessment of the data **relevance** the numerical values to be used as weights are defined as:

- Category 1: 1
- Category 2: 0.3
- Category 3: 0
- Category 4 was not considered

For assessment of the data **adequacy** the numerical values to be used as weights are defined as:

- *In vivo*: 1
- *in vitro* 0.3
- *in silico*: 0.1

A **final quality score** (weight) is then computed for each study as the average of the four scores of the quality criteria completeness, reliability, relevance and adequacy. This way, each study receives a single final quality score that can be used as weight of that study for further analysis.

For further details of the analysis and possible application of the results for hazard classification according to the CLP Regulation please refer to Basei *et al* (submitted).

3.3.4.3 Do you want to know more?

The following resources can provide more information:

Basei, G., Zabeo, A., Rasmussen, K., , Tsiliki, G., Hristozov, D. (accepted). Nanoimpact

Card, J., & Magnuson, B. (2010) A method to assess the quality of studies that examine the toxicity of engineered nanomaterials. Int. J. Toxicol., doi:10.1177/1091581810370720.

> *A 2-step method to assess the quality of nanotoxicity studies. The first step uses a publicly available tool to rank the reliability of the study based on adequacy of design and documentation of methods, materials, and results, providing a "study score." The second step determines the completeness of physicochemical characterization of the nanomaterial/nanomaterials assessed within the study, providing a "nanomaterial score."*

Comandella, D., Gottardo, S., Rio-Echevarria, I., Rauscher, H. (2020). Quality of physicochemical data on nanomaterials: an assessment of data completeness and variability in EU project databases, Nanoscale,12, 4695-4708, doi: 10.1039/c9nr08323e.

Examines the quality and completeness of data in the eNanoMapper. It found that many entries had missing information and this was attributed to a lack of harmonised data reporting and entry procedure

Fernández-Cruz, M., Hernández-Moreno, D, Catalán, J., Cross, R., Stockmann-Juvala, H., Cabellos, J., Lopes, V., Matzke, M., Ferraz, N., Izquierdo, J., Navas, J., Park, M., Svendsen, C., Janer, G. (2018). Quality evaluation of human and environmental toxicity studies performed with nanomaterials-the GUIDEnano approach. Environ. Sci. Nano. doi:10.1039/c7en00716g

An approach for a systematical and quantitative evaluation of the quality of environmental and human toxicity studies performed with nanomaterials. The approach builds upon previous initiatives and includes refinements to facilitate its application by users with limited toxicological expertise.

Gottardo, S., Ceccone, G., Freiberger, H., Gibson, P., Kellermeier, M., Ruggiero, E., Stolpe, B., Wacker, W., Rauscher, H. (2019) GRACIOUS data logging templates for the environmental, health and safety assessment of nanomaterials. https://publications.jrc.ec.europa.eu/repository/handle/JRC117733 doi:10.2760/142959.

The harmonised recording of experimental data on nanomaterial properties generated in different research projects is a key issue in nanosafety. This site is a store of templates that facilitate the reporting of data on endpoints. Each template relates the result of the measurement to the experimental conditions, protocols, method and instrument that have been used to generate it, thus ensuring reproducibility, comparability and re-use of the data by other scientists.

Magnusson, B., Örnemark, U. (eds.) (2014) Eurachem Guide: The Fitness for Purpose of Analytical Methods – A Laboratory Guide to Method Validation and Related Topics, (2nd ed.). ISBN 978-91-87461-59-0. Available from http://www.eurachem.org)

Examples of internationally accepted principles and procedures.

Schneider, K., Schwarz, M., Burkholder, I., Kopp-Schneider, A., Edler, L., Kinsner-Ovaskainen, A., Hartung, T., Hoffmann, S. (2009) 'ToxRTool', a new tool to assess the reliability of toxicological data. Toxicol. Lett., doi:10.1016/j.toxlet.2009.05.013

A software-based tool (ToxRTool) developed to provide comprehensive criteria and guidance for reliability evaluations of toxicological data. The tool aims to increase transparency and to harmonise approaches of reliability assessment.

3.3.5 Use of Similarity in the GRACIOUS Framework

Similarity of substances is a concept commonly used in REACH, during design and beyond. Screening large virtual libraries of molecular structures required the development of computational methods, allowing the user to go beyond an expert based qualitative similarity assessment. Similarities of non-NF chemicals are often derived from their chemical structures, which can be defined and assessed using a range of quantitative metrics, e.g. the AMBIT tool (see section 3.3.5.3). For non-NF chemicals different metrics are related to different biological effects, so the similarity assessment for some cases needs to be tailored to the purpose of the overall assessment and similarity of biological response can be introduced into the assessment. Within chemical regulation, different similarity methods have been recommended for read-across and have been applied to develop efficient screening methods to prioritize testing chemicals with a high potential of being hazardous.

By using methods to assess similarity we remove the need to provide strict thresholds or cut-offs. Instead, the similarity assessment could be thought of as generating a floating band which encompasses substances which are sufficiently similar to be grouped.

In REACH, similarity is explicitly requested for NFs for two distinct purposes:

a. For endpoint-specific grouping and read-across to generate the data required to determine hazard (REACH Annexes VII-XI). This is the purpose of the GRACIOUS Framework. Grouping of NFs to provide the data needed for a single hazard endpoint is also encouraged within REACH and the degree of similarity required to justify grouping can be defined by the assessor based on the results. However, any grouping and read-across decision must be based on both structural and biological mechanistic similarity.

b. To justify "sets of similar nanoforms"[1] during the registration step (REACH Annex VI). This is the purpose of e.g. the ECETOC Nano-App (Janer, 2021). Case studies used partially the same methods as requested by GRACIOUS IATAs (Janer, 2021a), but different assessment criteria are applied to justify which NFs belong in a single set, such that hazard, exposure and risk assessment can be performed jointly. As this measurement of similarity needs to apply across all hazard endpoints, the level of similarity needs to be high and the ECHA guidance (ECHA, 2019) establishes basic rules: All NFs must be of the same substance. Pristine and surface-functionalised NFs cannot belong in the same set.

Terminology

Quantitative similarity assessment is a specialised area of chemometrics and as such is associated with terminology that a user unfamiliar with this topic may find difficult to understand. Some useful terms are defined in Table 3.3.22.

[1] Please note that the GRACIOUS Framework does not support the generation of justification of a 'set of similar NFs'

Table 3.3.22: Subject specific technical terminology and their meaning

Term	Meaning
Property (intrinsic and extrinsic)	A property of a nanoform can be a basic physicochemical parameter (e.g. size, mass) required to identify a NF, or it can describe an aspect of the NF interaction with the immediate surroundings (e.g. reactivity, attachment efficiency). In the latter case the property depends on both the NF and its surroundings (extrinsic property), whereas in the former case the property is independent of the surroundings (intrinsic property).
(Scalar) descriptor	A single number, accompanied by units of measurement (e.g. nm). A scalar descriptor is the result of a reduction of a two-dimensional distribution of data points that characterises the data field in a way which is assumed sufficient for a specific purpose. Examples of scalar descriptors are D50 (median) for particle size distribution or LOAEL (Lowest-observed-adverse-effect level) for the dose response curve in inhalation toxicity.
Dynamic range	The range between the minimum and maximum value that a descriptor can have for a certain property. E.g. the mass-% content of an impurity ranges from 0% to 100%. For other properties, such as size, the dynamic range is unlimited.
Biologically relevant range	The range of descriptor values that has an impact on the biological behavior, such as for size, only values above 1nm are relevant.
Data matrix (matrix of data availability)	A matrix consisting of the group members/group candidates vs. corresponding set of available data on all relevant physicochemical, toxicological and ecotoxicological properties/endpoints for a specific IATA. The data matrix is the evidence base on which to formulate or decide a grouping or read-across decision. A data matrix contains all and only the evidence required by the IATA that applies to a specific hazard. Missing values are indicated by 'NA'. The matrix therefore helps highlighting the data gaps. The data matrix can be used to evaluate similarity between nanoforms for each hazard endpoint.
Pairwise similarity	Determined by application of some similarity algorithm to the data matrix entries of these two NFs. The resulting distance is also designated as "pairwise similarity score".
Similarity algorithm	A function that defines how far apart two data points are. Conventional examples are the Euclidean, Manhattan and Minkowski distances. The GRACIOUS white paper and case studies demonstrate additionally the x-fold algorithm, Baysian statistical algorithm, and arsinh algorithm. The x-fold algorithm is also used on the ECETOC NanoApp (https://www.ecetoc.org/tools/nanoapp/) to justify sets of nanoforms.
Distance or metric (a.k.a. similarity score)	The result of applying the similarity algorithm. A distance is a metric if it is nonnegative and symmetric, while the identity principle and the triangle inequality holds. The latter means that the distance between points A and B is less or equal to the sum of the distances between points A and C and between points B and C.
Multidimensional distances	The result of the application of some algorithm that assesses the data matrix of two NFs on several properties of the data matrix.
Data standardisation	In statistics standardised means that a data scaling transformation is applied per property to have variance 1 and mean 0.
Supervised and unsupervised machine learning methods, including cluster analysis	Machine learning algorithms generally can be divided into unsupervised or supervised. Regression and classification are supervised algorithms (because the training / fitting is supervised by the Y values). Clustering is unsupervised – clusters are identified solely by X data, without taking into account any Y data. Examples are hierarchical and non-hierarchical unsupervised algorithms such as Hierarchical Cluster Analysis (HCA), k-means algorithm, Density-Based Spatial Clustering (DBSCAN), spectral clustering. Dimensionality can also be reduced by other methods (e.g. Principal Component Analysis (PCA)),
Benchmark materials and Representative Test Materials (RTM)	All materials used in GRACIOUS as benchmark or reference materials are representative test materials in the metrological sense. They serve as a point of reference to support the interpretation and assessment of results obtained on a new test material. A representative test material is a

	material from a single batch, which is sufficiently homogeneous and stable with respect to one or more specified properties, and which implicitly is assumed to be fit for its intended use in the development of test methods which target properties other than the properties for which homogeneity and stability have been demonstrated. RTMs used in GRACIOUS are well-characterised nanomaterials, e.g. from the JRC repository (https://ec.europa.eu/jrc/en/scientific-tool/jrc-nanomaterials-repository). For some assays, they also serve as positive and negative controls, but controls could also be non-particulate chemicals.
Fingerprint/fingerprinting	A unique set of descriptors indicating the presence of particular functionalities in or on a NF, as based on specialized analytic techniques.

3.3.5.1 Instructions to assess similarity when justifying grouping in the GRACIOUS Framework

The following section describes the steps recommended for a user to quantitatively measure similarity to justify a grouping decision. As stated previously this can be done for a single DN or across the entire data matrix. These two situations are described separately.

Similarity within a single Decision Node

1. **If not recommended specifically by the pre-defined IATA, identify whether similarity is the best method to reach a conclusion for the Decision Node of interest**

The user should assess each DN within an IATA to decide whether a cut-off or similarity/floating band approach is most appropriate. The cut-off approach is most easily assessed, but as biological responses often occur along a continuum, the justification of selecting a particular value might be difficult. Candidate NFs may give results that fall either side of the cut-off value, so although the results are similar, the use of the cut-off could put them in separate groups. This should lead a user to either re-assess the cut-off value, to move to a higher tier (where the issue may persist) or to use a similarity/floating band approach instead. The approach to assess similarity could be governed by the purpose for grouping. Some regulations have established definitions for categories of substances, so if the purpose for grouping is regulatory it may be useful to use these cut-off values. Where relevant, such regulatory cut-off values are often already included in the pre-defined IATAs. As an example, a WHO (World Health Organisation) fibre is described as a sample containing > 0.1 % of inhalable particles being > 5 µm in length and with an aspect ratio of > 3:1. If the use of the inhalation IATA for hypothesis **H-I-1** is for regulatory purposes, using these cut-off values is justified.

Where no obvious cut-off value is apparent, the similarity/floating band approach is recommended (Figure 3.3.12). The Framework requires that similarity across all DNs is used to reach the final grouping decision.

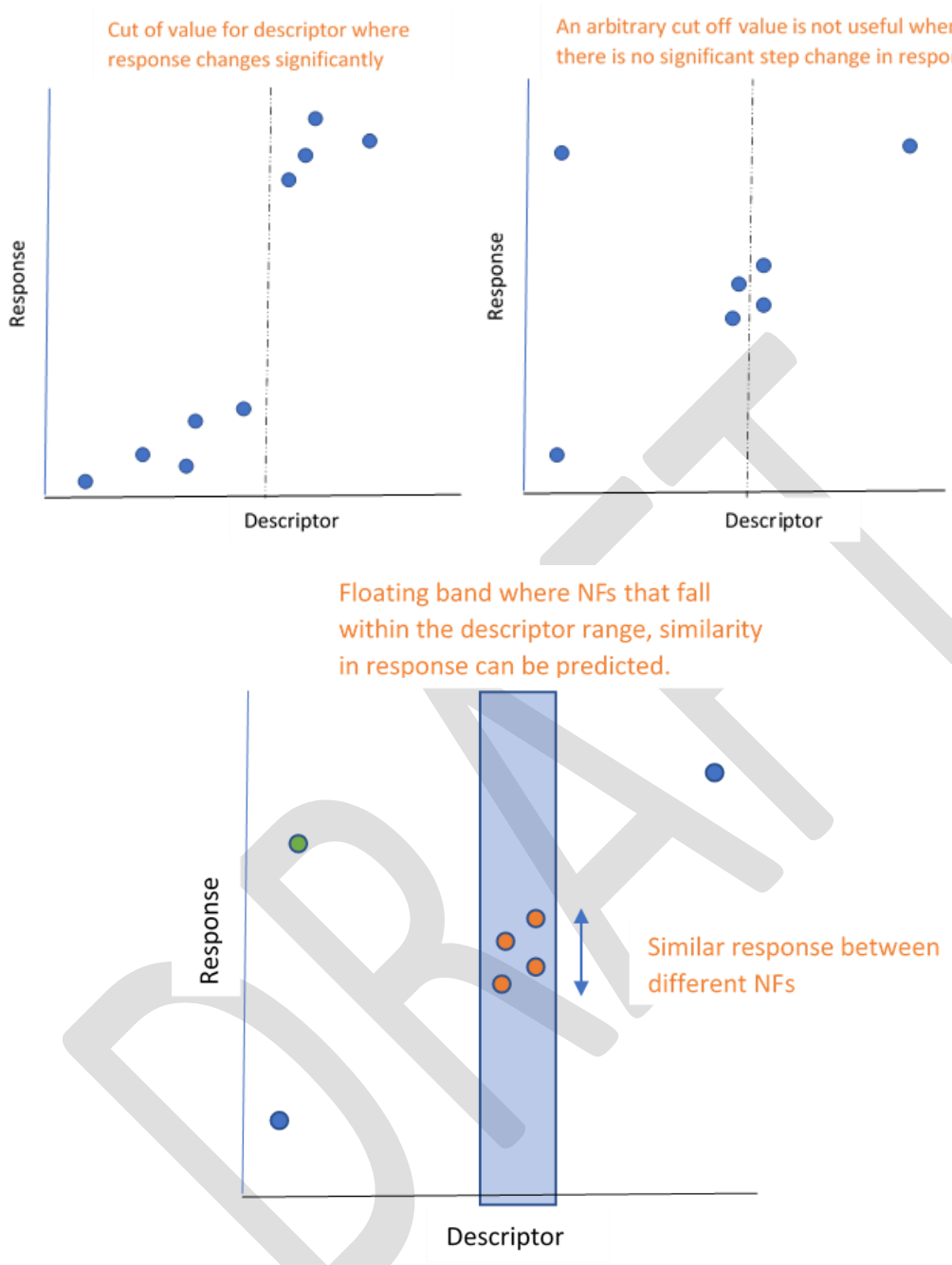


Figure 3.3.12 Choosing how to define similarity: i) A step change in response at a specific value allows a cut-off to be used (e.g. HARN IATA); ii) where there is no step change, a cut-off value is not appropriate; iii) instead defining a floating band that describes a range of descriptor values that will give a similar response may be more useful (it should be noted that in this example, the blue and orange NFs may be showing a linear relationship between descriptor value and response. If this can be proven with other experiments, it might be possible to group blue and orange NFs. The green NF appears to be an outlier, if it can be shown that this elicits a response via a different MoA, this strengthens the grouping hypothesis for the other NFs).

2. Assess the dynamic and applicability range of the property under examination

The full dynamic range of a property might not be relevant to measure, either due to lack of biological relevance (e.g. particles over a certain size cannot be inhaled), or due to a lack of method accuracy. Also regulations may limit the relevant range, e.g. diameters below 1 nm are not relevant. The applicability range describes the range of values within which the property can be measured reliably for the members of the group. The user should choose a method to assess similarity that allows them to distinguish between property values that fall within and outside the applicability range (Figure 3.3.13). Outside of the biologically relevant range, any pairwise comparison of NFs should assess them as sufficiently similar, because the biological activity is not impacted by this property in this range, e.g. the NanoFASE project has shown that the environmental transport is not impacted by attachment affinity values outside of a certain relevant range (Meesters *et al*., 2019).

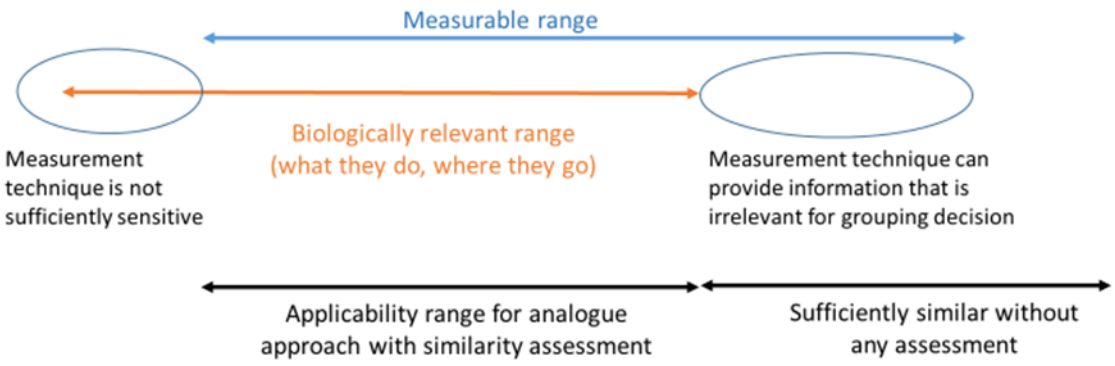


Figure 3.3.13: An applicable range for the assessment of similarity depends both on biological relevance and the limits of the chosen analytical technique.

3. Consider data reduction to scalar descriptors

Data is often available as a distribution (e.g. a concentration response curve for hazard) which is too complex to use routinely for a similarity assessment with large data sets. Novel algorithms such as the Baysian statistical approach can generate a pairwise similarity score from distributions, but are not a routine tool yet. Instead, data distributions can be converted to a single value known as a scalar descriptor (e.g. LC50). It should be noted that different scalar descriptors can be derived from the same set of data (e.g. T25 or BMD10 for carcinogens). For a similarity assessment the same scalar descriptor would be needed for each candidate NF. However, in some cases this may oversimplify the results, as schematically exemplified on dose-response relationships (see Figure 3.3.14). To ensure that no important information is lost in data reduction, one can a) perform the data reduction both in mass dose metric and in surface area dose metric, to check for consistent scaling by surface area, or b) check qualitatively the shape of distributions to select a descriptor that is sensitive to the differences (see Figure 23), or c) use one of the novel algorithms on two-dimensional data. Strategies a) and b) contribute to the demonstration of a common MoA, which is essential to justify grouping.

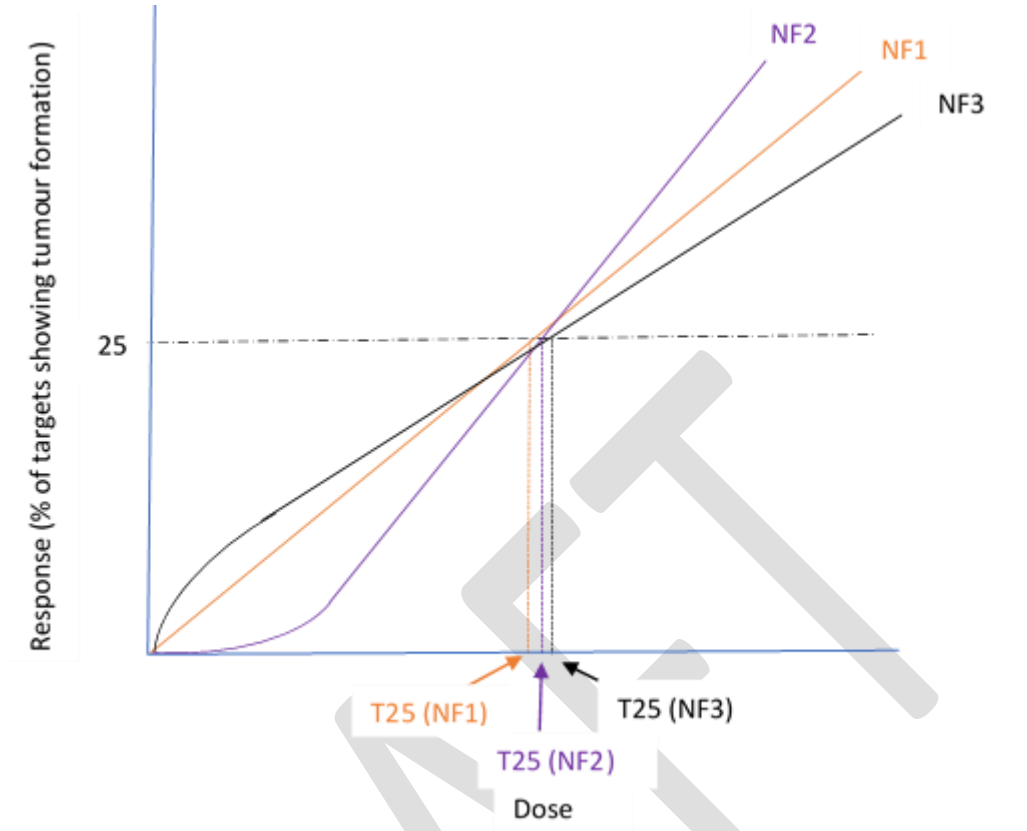


Figure 3.3.14: The dose response curves from a carcinogenicity study for three NFs. If the data were reduced to a single dose descriptor, T25, the NFs would appear similar. Examination of the full dose-response curve shows the NF1 gives a linear response, NF2 a sublinear response and NF3 a supralinear response. This should indicate the potential for different MoA and hence no similarity between the NFs.

4. Understand achievable data accuracy

When assigning a cut-off value or the range of a floating band, the impact of accuracy and reproducibility must be taken into account (also see Quality section 3.3.4). Reproducibility checks on basic physicochemical information, as required to register a NF in REACH, have shown that these measurements can be reproduced across four experienced laboratories with just a few % accuracy limits. This is important to determine because ECHA guidance requires "well-defined boundaries of the group", after the successful demonstration of similarity. However, similarity assessments are based on IATA DNs that typically use interaction properties, such as reactivity or dissolution rates in specific media. The draft OECD test guidelines and methods for these properties are only starting to be tested with interlaboratory comparisons, and are often limited to a 1.5-fold to 2-fold reproducibility. Available data from databases should also include this information (if possible) and if it is not available its use may need to be reconsidered.

5. Decide the type of similarity assessment required

For regulatory purposes of the GRACIOUS Framework, a property-by-property evaluation of the data matrix is recommended to generate a pairwise similarity assessment, i.e. each NF is assessed against every other NF of the candidate group. The results can be plotted on a triangular similarity matrix for each property. This means that for one grouping hypothesis, all DNs of the associated IATA need to be assessed individually using the property-by-property evaluation. If any NF for any DN is not found to be sufficiently similar for that property, then that NF should be considered for removal from the group. If for any DN there is no evidence of similarity between the NFs then the whole hypothesis should be rejected.

Multidimensional distance metrics, for example dendrogram clustering approaches, were found to offer unexpected insights into the overall similarity of very different materials, but it is a major challenge to select a distance metric that is appropriate for all dimensions (i.e. all properties), and inappropriate data transformation can lead to false conclusions. The multidimensional tools are therefore currently difficult to use in a regulatory context. If materials are identified as less similar when using these methods, the user may need to additionally consider their ranking in individual properties, because rankings are not represented by distances, but may be important to justify read-across. When used for exploratory scientific purposes, the robustness should be challenged by carefully selecting the distance metric, and by comparing to other defendable distance metrics. The multidimensional approaches are not generally recommended for regulatory purposes, instead they are primarily tools of discovery.

6. Choose the method for quantifying similarity and apply it to the data sets

There are a wide range of different mathematical approaches to quantifying similarity. Some are more appropriate for pairwise assessment and others for cluster analysis, so the user needs to match the method with the type of assessment decided in the previous step. Each has their strengths and weaknesses, so choosing the best approach will be guided by a range of factors including

- Purpose for Grouping
- Nature of study(s) being assessed
    - Single point result or dose response relationship (curve)
- Type of relationship being assessed
    - One-to-one relationship
    - Category
- Assessment of a single study or across all DNs

Some examples of quantitative similarity methods used in the development of the GRACIOUS Framework are shown in Table 3.3.23.

Table 3.3.23: Quantitative methods to measure similarity suitable for use within the GRACIOUS Framework.

Method	What does it do?	Strengths	Weaknesses
Euclidean distance	It describes the length of the line segment between two points. For scalar descriptors it is equal to the absolute value of the difference between the scalar values. Equations exist for multidimensional metrics.	Standard method, easy to implement, multi-dimensional. Interfaced to the GRACIOUS blueprint via the eNanomapper database.	Assumes data follows normal distribution; does not work with missing data.
Cluster analysis	Cluster analysis aims to discover two-dimensional patterns in the data matrix, searching for similarities between NFs and properties. NFs clustered together (i.e. grouped) are considered to be more similar between one another, compared to all other NFs belonging to other clusters.	Many clustering methods available, easy to use, visualisation possible.	Does not work with missing data.
X-fold comparison	When comparing descriptor values for two different NFs, the x-fold comparison divides the smaller of two values by the larger.	Simple. Integrated in the GRACIOUS blueprint (incl. data cropping to the relevant range)	Does not correct for differences in the biological relevance over the dynamic range, such as noise and accuracy limits, but data can be cropped to the relevant range.
Arsinh OWA model	Cluster analysis that applies a transformation that allows introduction of new members.	Based on absolute distance metric, derived groups are not relative to the assessed entities.	Requires establishing a proper threshold for scaling.
Bayesian model	Compares two sets of values using nested sampling. Standardized data are compared to determine whether they are derived from the same normal distribution.	Able to incorporate literature or previous knowledge from public data.	Difficult to use as it needs adjustment for different statistical distributions depending on the data analyzed.

7. Identify the similarity value that will define similar NFs for the study under assessment

If a pairwise similarity assessment is done, the user will now have a range of quantitative values of similarity between each pair of candidate NFs. The user now needs to decide which values correspond to the NFs being similar and which to the NFs being dissimilar. It is highly recommended to demonstrate the biologically relevant range by inclusion of representative test materials (from the JRC repository) in the same assessment. This demonstrates sensitivity of the method and algorithm. Hence overall materials of three substances (two RTMs and the NFs of the candidate group) are included in the assessment. There are a number of factors that must be considered.

a. Sensitivity of MoA to the variable being measured in the study

b. Applicable and dynamic range of the study

c. Accuracy and reproducibility of the study

d. Purpose for grouping

If the purpose for grouping is to inform SbD decisions, the degree of similarity required to confidently assign two NFs as being similar is less strict than for regulatory obligations.

The pairwise similarity can be visualised across all candidate NFs using a triangular similarity matrix (Figure 3.3.15).

NF1		30	7	13	12	1000	2500
NF2			125	7000	79	1500	400
NF3				6	9	850	3000
NF4					1.2	5000	300
NF5						6500	1250
RTM1							
RTM2							
	NF1	NF2	NF3	NF4	NF5	RTM1	RTM2

Similarity	Range of x-fold difference
Similar	< 15
Undecided	15 - 100
Not similar	> 100

Figure 3.3.15: A hypothetical example of a pairwise similarity assessment of the reactivity of 5 NFs. The x-fold difference between NFs is given in each box and the colour indicates whether this is regarded as similar (there is no scientific justification for the limits of the similarity categories, the numbers are chosen to exemplify the process). The method used to quantify similarity was the x-fold method and any values within 15 x of each other were defined as similar and any values > 100 x different were assessed as definitely not similar. A decision could not be reached for NF-pairs whose reactivity differs 15 – 100-fold and further experiments would be needed (higher tier). NF2 appears not to be similar to the other candidate NFs, so it might be necessary to exclude it from the potential group.

Based on the similarity scores of RTMs and orientating case studies, we concluded that the x-fold, Bayesian and Arsinh-OWA distance algorithms are mutually consistent in scoring NF pairs. The very popular Euclidean distance is also useful, but only with Yeo-Johnson data transformation which enhances consistency with the other algorithms, albeit not perfectly. The tier 1 score of a NF pair with known tier 3 similarity can be indicatively set at or below 1.3 (Yeo-Johnson Euclidean) and at or above 1.5 (Bayesian). For the x-fold metric acceptable similarity can be indicatively set at or below 5-fold for many properties, whereas the comparison of opposite controls (i.e. the pair of representative test materials, e.g. ZnO NM110 representing quick dissolution and TiO_2 NM105 representing very slow dissolution) scores between 100-fold to 1000-fold.

8. **Identify how to use the similarity assessment.**

How the user proceeds through the IATA once similarity within a study has been determined will depend on a number of factors, including:

- What was the result of the similarity assessment? If there are some NFs where similarity cannot be determined, it may be necessary to move to a higher tier study within the DN (for all NFs). If one or more NFs are not sufficiently similar to other group members with respect to one or more DNs, they may need to be removed from the group.
- The purpose of grouping. As with step 7, a regulatory purpose may require a greater degree of confidence in the similarity between the candidate NFs than other purposes.

Quantitative similarity assessment across the whole data matrix

The steps required for this analysis are closely related to those for a specific DN, but as the results have already been processed during the IATA, it is possible to start at Step 5.

1. **Decide the type of similarity assessment required**

The final grouping decision is made based on an assessment of the whole data matrix for each candidate NF still under assessment. The data matrix contains all and only the properties requested by the specific IATA.

2. **Choose the method for quantifying similarity and apply it to the data sets**
3. **Identify the similarity value across all NFs and DNs that will be used to decide whether a group can be justified**

For regulatory grouping purposes in particular, grouping needs to be scientifically justified. Therefore, the choice of measurement of similarity needs to be supported by the science and expert judgement. The inclusion of the simplest approach would be to include all DN pairwise similarity assessments into a single triangular similarity matrix, where similarity between NFs can only be ascribed if there is similarity between the NFs in all DNs.

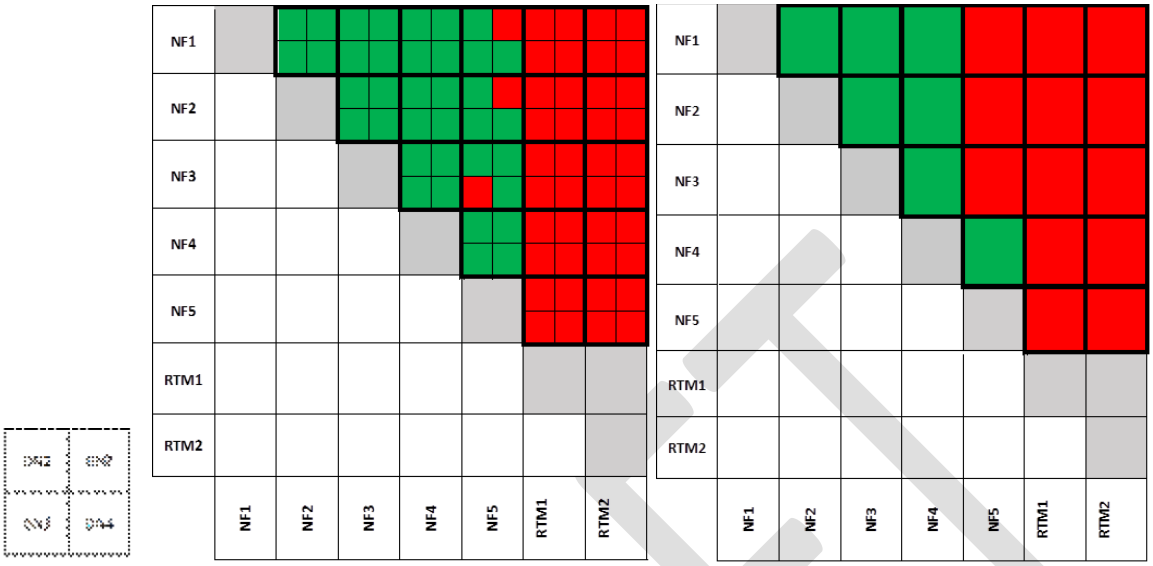


Figure 3.3.16: An example of combining pairwise similarity assessments from all DNs to make a grouping decision across the whole IATA (Green = similarity demonstrated; Red = similarity not demonstrated). I) All pairwise property-by-property similarity assessments between each pair of NFs are shown across four DNs. II) Overall similarity between NFs across all DNs, based on the assumption that all DNs must be sufficiently similar for a pair of NFs to be sufficiently similar. To demonstrate the sensitivity of the methodology, the pair of RTMs (i.e. the opposite controls for each property) should always result in a non-grouping decision for each property. For simplicity, we plotted here the case that RTM1 and RTM2 serve as controls for all four DNs, but in reality, each DN may have its own pair of RTMs to span the biologically relevant range for that property.

In the example shown in Figure 3.3.16, the final group would be NF1, NF,2, NF3 and NF4. Even though NF4 and NF5 are similar, grouping requires that all NFs are similar to each other, so NF5 cannot be included in the group.

4. Identify what will be the impact of the grouping conclusion

Once a group has been identified and justified, the user will then need to decide what to do with the group. This decision will refer directly back to the initial purpose of the grouping (see section 3.2.1).

3.3.5.2 Worked Example: Quantitative assessment of similarity of nano and non-nanoforms of organic pigments.

Context: Three samples of an organic pigment, DPP nano, DPP non nano and DPP pre-mixed are potential products for development. The manufacturer wishes to understand whether read-across from one form to the other for hazard endpoints might be possible. The endpoints are believed to have an inflammatory MoA. One of the tiers of the IATA includes *in vitro* inflammatory studies. In this worked example, one inflammatory assay, Nrf-2 activation is discussed, but the IATA recommended to use three other assays. For more details on these studies please see Ag Seleci (2021).

1. If not recommended specifically by the pre-defined IATA, identify whether similarity is the best method to reach a conclusion for the Decision Node of interest

Current understanding of the MoA does not indicate a threshold where inflammation is triggered, so similarity is the best approach to support a grouping conclusion.

2. Assess the dynamic and applicability range of the property under examination

RTMs were used to set the biologically relevant range for the inflammatory studies. Manganese oxide was used as the positive control and barium sulfate was used as the negative control as their respective effects on inflammatory response were well understood. The responses to the RTMs were used as the upper and lower limit of the applicability range for this worked example. It is acceptable for the test samples could lie outside the range of the RTMs providing that they fall within an order of magnitude of one of the RTMs.

3. Consider data reduction to scalar descriptors

Reduction to scalar descriptors was not deemed appropriate in this situation because it was thought assessment of the similarity of concentration-response curves could give a more robust scientific justification for grouping. This decision was driven by obvious differences in curve shape and the magnitude of response over a wide range of exposure concentrations.

4. Understand achievable data accuracy

To understand the achievable data accuracy, graphs were generated for all organic pigment sample values to check for any discrepancies or inconsistencies based on experts' opinion. Plotting concentration-reactivity curves was deemed important in this case to compare samples with RTMs and validate their effects on the inflammatory response.

5. Decide the type of similarity assessment required

Pairwise similarity analysis was performed in a 3-step manner employing three different criteria:

- Similarities between shapes of reactivity concentration-response curves.
- Similarities between the concentration ranges.
- Similarities between the reactivity ranges.

This was done because it was felt this would give the strongest scientific justification for the similarity conclusion (e.g., two dose-response curves can display a very similar shape across two very different response ranges, so simply comparing the shape was not considered sufficient in this scenario).

6. Choose the method for quantifying similarity and apply it to the data sets

In this scenario, all three criteria were considered to account for the concentration-reactivity curves similarity, the concentration range similarity and the reactivity range similarity.

7. Identify the similarity value that will define similar NFs for the study under assessment

The similarity of the curves was measured via Bayes Factor calculations. Similarities between ranges concentration and reactivity data were quantified using the Manhattan distance metric in both cases. This was used additionally to the Bayes Factor calculations to cope with large differences in the concentration ranges measured.

The final similarity score reported was a weighted average distance metric, which for each pair of NFs, combined the BF value with quantification of the distance between the ranges of the response reactivity values d_R and the distance between the ranges of their concentration d_D.

$$similarity\ score = (0.3 * BF) + (0.5 * d_R) + (0.2 * d_D)$$

The pairwise similarity scores were rescaled to the range (0,1) and plotted onto a triangular similarity matrix (Figure 3.3.17). In this example, an arbitrary value to distinguish similar and dissimilar sample pairs was not defined because the results were intended to be used in conjunction with other assays to draw a final conclusion. Colour-coding was used to differentiate different degrees of similarity where highly similar pairs of NFs are shown here with red colour.

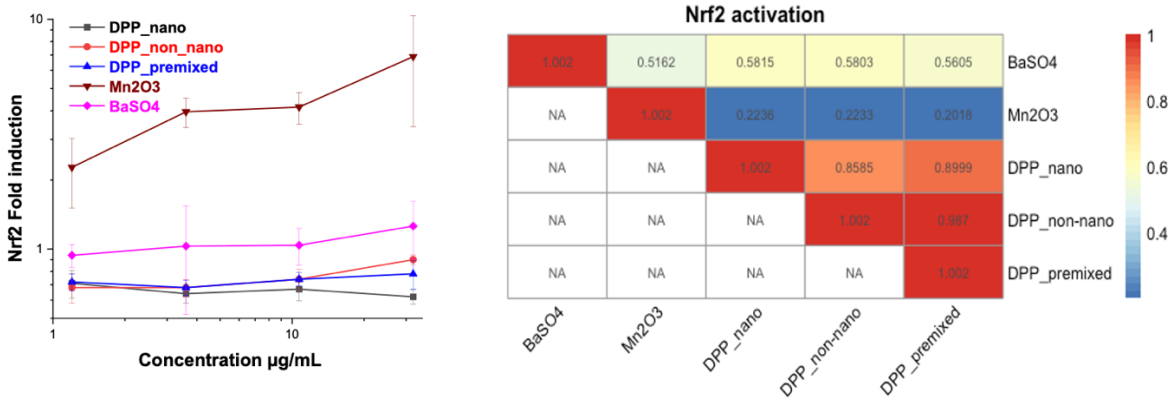


Figure 3.3.17 Reactivity-concentration curves (left hand-side) and the corresponding pairwise similarity matrix (right hand-side) of three organic pigment samples and two RTMs for the Nrf2 activation assay.

These results showed a high degree of similarity between DPP non-nano and DPP premixed and a reasonable degree of similarity between both of these bulk forms and DPP nano. All three samples showed some similarity to the negative control RTM, barium sulfate and all most none to the positive RTM, manganese oxide.

8. Identify how to use the similarity assessment.

As previously stated, this worked example was one of four inflammatory assays examined, so the results presented were not used in isolation to reach a conclusion. However, the results of the other assays confirmed that each form of the pigment were similar and that the results of this DN would support the grouping of these forms when the whole IATA was assessed for similarity.

(Note. *In vivo* inhalation studies were available for both the nano and non-nano forms and both had NOAEC of > 30 mg m^{-3}. The worked example would support the use of read-across to these values for the pre-mixed sample, if the results of the remainder of the IATA also support the grouping of these samples).

3.3.5.3 Do you want to know more?

The following resources can provide more information:

AMBIT (http://cefic-lri.org/toolbox/ambit/)

> *The AMBIT system consists of a database including more than 450.000 chemical structures and REACH dataset of 14.570 substances. AMBIT contributes to the safer use of chemicals and a reduction in testing and innovation cost by making it easier for companies to comply with regulations governing chemicals. Users can search and access a wide range of existing information and prediction about a chemical. This process makes the tool both unique and powerful, particularly for data-poor small and medium sized enterprises (SMEs).*

ECETOC Nanoapp. Available at https://www.ecetoc.org/tools/nanoapp/

> *ECETOC's NanoApp is a tool designed to define the boundaries of sets of similar NFs and to generate a justification for the REACH registration. It must be noted that it is not intended to group NFs for the purposes of addressing specific regulatory endpoints or hazard concerns.*

ECHA (2019). Appendix for nanoforms applicable to the Guidance on Registration and Substance Identification.

Available at https://echa.europa.eu/documents/10162/13655/how_to_register_nano_en.pdf/f8c046ec-f60b-4349-492b-e915fd9e3ca0

> *The guidance document gives instructions for registrants to register nanoforms of substances under REACH. The concept of sets of similar nanoforms is introduced and guidance to establishing and justifying them is given.*

Floris, M. and Olla S., (2018). Molecular Similarity in Computational Toxicology. Methods in molecular biology (Clifton, N.J.), 1800: p. 171-179.

> *Recommends similarity methods to be used for read across*

Janer G., Landsiedel R., Wohlleben W. (2021). Rationale and decision rules behind the ECETOC NanoApp to support registration of sets of similar nanoforms within REACH. Nanotoxicology 15.2 (2021): 145-166, doi.org/10.1080/17435390.2020.1842933.

Explains the development of the ECETOC NanoApp tool that uses pairwise similarity to help identify sets of similar nanoforms for REACH registration.

Janer et al. (2021a, in print). Creating sets of similar nanoforms with the ECETOC NanoApp: real-life case studies. Nanotoxicology

Examines case studies that use the ECETOC NanoApp to create and justify sets of similar nanoforms

JRC Nanomaterials Repository. Available at https://ec.europa.eu/jrc/en/scientific-tool/jrc-nanomaterials-repository

The JRC hosts a repository of representative industrial nanomaterials (NM) including nanomaterials studied in the OECD testing programme and large research projects. Each type of material in the repository has been sourced as a large single batch which has been sub-sampled into individual vials to produce the first collection of thoroughly characterised nanomaterials available for benchmarking in research and regulatory studies.

Meesters, J., Peijnenburg W., Hendriks A., Van de Meenta D., Quik J. (2019). A model sensitivity analysis to determine the most important physicochemical properties driving environmental fate and exposure of engineered nanoparticles. Environ. Sci.: Nano, 2019. 6(7): p. 2049-2060, doi.org/10.1039/C9EN00117D.

Investigates and identifies key characteristics that drive the environmental fate of nanoparticles.

Mellor C., Marchese Robinson R., Benigni R., Ebbrell D., Enoch S., Firman J., Madden J., Pawar G., Yang C., Cronin M. (2019)., Molecular fingerprint-derived similarity measures for toxicological read-across: Recommendations for optimal use. Regulatory Toxicology and Pharmacology, 101: p. 121-134.

Recommends similarity methods suitable to be used for read across

Wassenaar, P., Rorije E., Vijver M., Peijnenburg W. (2021). Evaluating chemical similarity as a measure to identify potential substances of very high concern. Regulatory Toxicology and Pharmacology, 119: p. 104834-104834, doi.org/10.1016/j.yrtph.2020.104834.

Recommends methods to use similarity to prioritise chemical substances by their potential for hazardous properties

3.4 Outcomes from Grouping – Read-across

Once a group of NFs has been identified using the Framework, the user can conduct read-across from one or more source NFs or non-NFs, for which data and information exist, to a similar target NF where information is lacking. To allow this, read-across requires development of a robust scientific explanation of similarity between the source(s) and the target for that specific endpoint. Once data gaps are filled for all group members, the user can apply these to the original purpose of grouping.

Read-across is typically used in a regulatory setting, although the GRACIOUS Framework supports its use for other applications such as SbD, to inform the need for additional studies, risk management measures, or communication of potential safety issues along the nano-enabled product value chain. For SbD, the required level of detail of information on safety increases as the *Stages* of product development progress. When read-across can be applied for one or multiple endpoints, this may lead to more cost-efficient gathering of information for regulatory registration before market launch. It may be possible therefore to anticipate the regulatory application of read-across early in the development progress, so that less resources may be needed for information gathering for subsequent regulatory approval. The read-across approach supported by the GRACIOUS Framework follows the process outlined by ECHA.

Different types of read-across

In their guidance, ECHA highlights that a number of different types of read-across approaches are available to a registrant, depending on the number of substances in a group and how they are related within the group.

Analogue approach: Grouping based on a very limited number of chemicals (possibly only two), where trends in properties are not apparent.

Type of read-across relevant to the analogue approach

- One-to-one (one analogue used to make an estimation for a single chemical)

Category approach: A group of chemicals whose physicochemical and human health and/or environmental toxicological properties and/or environmental fate properties are likely to be similar or follow a regular pattern as a result of structural similarity (or other similarity characteristic)

Types of read-across relevant to the category approach

- Many-to-one (two or more analogues used to make an estimation for a single chemical)
- One-to-many (one analogue used to make estimations for two or more chemicals)
- Many-to-many (two or more analogues used to make estimations for two or more chemicals)

The GRACIOUS project has largely identified case studies that address the analogue approach, but there is also at least one example where sufficient data exists to support use of the category approach.

3.4.1 Instructions for the application of read-across to a group of nanoforms

The read-across methodology includes (i) identification of **source materials**, (ii) generation and justification of a **read-across hypothesis** to provide the scientific basis for using the data from a source substance/NF to fill a data gap for a target NF, and (iii) application of **methods to fill the data gaps** (Figure 1).

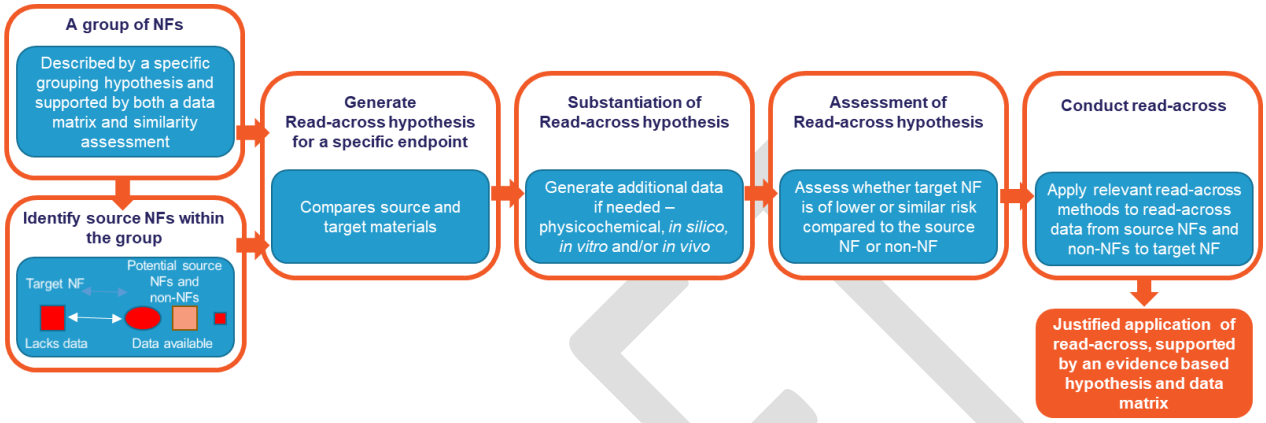


Figure 3.4.1: Generation and justification of a read-across hypothesis to provide the scientific basis for using the data from a source substance/NF to fill a data gap for a target NF. Adapted from Stone et al (2020).

1. **Review purpose and data matrix to identify data gaps that need to be filled by read-across**

It is essential that the read-across focuses on the data-gaps most applicable to the original purpose of grouping. For example, if the purpose of grouping was to provide the *in vivo* data needed to address regulatory endpoints for all group members, examination of the data matrix generated during the IATA will identify which group members have the required experimental data and which do not.

2. **Generate a read-across hypothesis**

For read-across, the same grouping hypothesis can be edited to focus on filling specific data gaps for a specific endpoint. The read-across hypothesis will thereby address whether the target NF is likely to have a similar hazard or whether it is less hazardous than the source NF(s) or non-NF(s). The data matrix generated for grouping can be used to double check that the read-across hypothesis is appropriate (i.e. that the target material is similar enough to or less hazardous than the source to be grouped together and share data for one or more endpoints).

A read-across hypothesis will be very closely related to the grouping hypothesis tested by the IATA. It will be specific to the user's purpose for entering the GRACIOUS Framework. A useful template to follow would be

"Endpoint will be provided for all group members by reading across data from ***Source NF*** to ***Target NFs***. This is justified because the ***Similarity Assessment*** has confirmed that ***Grouping Hypothesis Title*** is accepted."

3. **Substantiation – Check proposed source NF or non-NF is part of the group and generate further data if required.**

Either at the start of the Framework or at the completion of the data matrix with existing information a potential source NF or non-NF should have been identified. The user should now check whether this source NF or non-NF is still part of the group. If this is the case, the user can move to 'Assess read-across hypothesis' - below. If it is not, then the user has two options:

a) Identify a new NF from a database to be included in the group

The user should be aware that choosing a new NF to act as a source for the read-across will require that all DNs in the IATA are completed for this NF. If the data gaps cannot be completed with existing information, the studies in the IATA would need to be performed on this new NF. The user must be aware that it may be problematic to identify this NF to conduct the required studies.

b) Select NF from the group to fill data gap experimentally

If it is not practical to use existing data on a new NF, the user will need to identify one group member to become the source NF. The hazard endpoint study (usually a tier 3 method) will then need to be commissioned with this NF. It is recommended that the NF selected should be the one judged to be likely to give the most adverse outcome (preferably the selection should be substantiated/justified).

4. **Assess read-across hypothesis**

As the read-across hypothesis is based on the grouping hypothesis, it can be assumed that similarity of group members has been measured and confirmed. In this step the user needs to ensure the method used to read-across from the source to the target NFs is appropriate. This will depend on the type of read-across approach proposed.

- Analogue approach (One-to-one) – Source NF is likely to demonstrate almost identical or marginally more hazardous behavior than target NF.
- Category (Many-to-one) – If the source NFs hazard lies across a range, the target NF should be predicted to fall within this range so a processing value can be justified.
- Category (One-to-many) – Source NF is likely to demonstrate almost identical or marginally more hazardous than target NFs.
- Category (Many-to-many) - If the source NFs hazard lies across a range, the target NFs should be predicted to fall within this range as far as practicable, as greater confidence can be applied to interpolation than extrapolation

5. **Fill data gaps for target NFs**

Methods to fill data gaps can either be qualitative (based upon expert judgement) or quantitative. When applying quantitative read-across, there are four general ways of estimating the missing data point (listed below). The choice of method is dependent upon purpose, and currently these are suggestions rather than rules.

- **Most conservative value**

The data gap is filled by copying the most conservative value of the closest analogues or the most conservative value in the (sub)category. The conservative approach should be the default method for filling data gaps. Justification would need to be provided if this method is not used. Examples of scenarios for which this approach should be prioritised include:

  - Cases where there is a specific concern (e.g. if the NF is likely to be bio accumulative).
  - Cases which include DNEL or PNEC values.
  - Cases where there is uncertainty (e.g. with 2 small data sets) and therefore a lack of confidence.

This approach is applicable to all read-across approaches within both the analogue and category approach. It does require the source NF or non-NF to be the group member that can be predicted to display the most severe hazard properties (e.g. the most conservative value).

- **Copying from one source NF or non-NF**

The endpoint value of a source chemical can be simply copied and pasted into the relevant empty sections of the data matrix. For example this could be from the closest analogue in a (sub)category. Often the data copied would be a range including confidence intervals.

If more than one study is available for the same source material, the user needs to report all of the studies conducted, the results of these studies and an assessment of the quality of each data set (see section 3.3.4). Obviously more weight should be given to studies of higher quality. If there is more than one study of high quality then the user should employ the most conservative. If the most conservative data is not used then a justification of why that data is not used would be required. ECHA provides extensive guidance on how this should be performed for all substances (https://echa.europa.eu/documents/10162/13632/information_requirements_r8_en.pdf/e153243a-03f0-44c5-8808-88af66223258).

This approach can be used when there is only one source NF or non-NF, or where there is a high degree of similarity between group members. It may be used in the one-to-one and one-to-many read-across approaches. When the source NF can be predicted to display the most severe hazard properties, this method is identical to the most conservative method.

- **Processing values from multiple source NFs or non-NFs**

Processing includes manipulations such as calculation of the average, a weighted average, or median value, in order to use the most representative value. This approach clearly requires endpoint values from two or more source NFs or non-NFs, and all data for all source NFs or non-NFs would need to be inserted in the data matrix for inclusion in a subsequent dossier for regulatory purposes. If the user possesses a large data set (e.g. > 10 values that are close), then averaging will be useful. However, approaches such as the weighted average are helpful where there is large variation in the data set (small or large) as this approach reduces the contribution of the most distant data points.

Instead of a single value, a data range could be generated by this approach, along with confidence intervals if they are available. Again an assessment of the quality of each data set would be needed, with more weight given to studies of higher quality. Finally, with this type of approach it is important to check

that the design of the studies employed are in alignment with the read-across hypothesis. Finally, this method could be prioritised over the closest analogue approach if a lot of data exists, although this also depends upon the degree of similarity between the source and target NFs. It may be more appropriate where the most conservative value has been calculated using a lower quality study (see Quality section).

This method is appropriate for the category read-across approaches many-to-one and many-to-many but will need strong justification to be used instead of the most conservative value method.

- **Modelling approaches**

An internal Quantitative Structure Activity Relationship (QSAR) can be used to scale the available experimental results from two or more source NFs or non-NFs to the target NF. This approach is more suitable for a category approach where more data is available. At this time few QSARs are available for NFs that are sufficiently well developed and tested to support read-across for a variety of data types and read-across hypotheses. For regulatory purposes a validated model should be used, but for SbD and precautionary risk management purposes the users are free to use models as they see fit, although they need to be aware of the lower confidence arising from using an unvalidated model.

This method is applicable to the category approaches many-to-one and many-to-many and may be useful where the similarity between group members is lower but a trend across the group can be seen. For the purpose of the GRACIOUS Framework read-across application, the conservative approach is listed first, as this should be the default for regulatory read-across of NFs. The closest analogue and the processed endpoint (e.g. average) are listed equally, with the choice depending upon whether you have enough data to average or not. Tools for generating an internal QSAR have not really been developed and so this approach is listed last.

TIP for new users

Read-across is a regulatory method and as such a user should follow any rules stated in the relevant regulation if the purpose of grouping is to satisfy regulatory endpoints. If a different purpose of grouping is applicable, the user can relax these rules, for example

- *Read across between NFs and non-NFs made of different substances.*
- *Work with higher levels of uncertainty (either lower tiers of testing, or data sets of lower quality)*
- *Use a wider range of read-across methodologies, including qualitative methods.*

6. Document read-across justification

The level of documentation required to justify the read-across will depend on the purpose of grouping. Regulations such as REACH require read-across justification to be clearly and comprehensively documented and submitted with the regulatory application to allow regulators to assess the NFs with confidence. Grouping for SbD or precautionary risk management purposes does not require such formal documentation, but some recording of the read-across justification would be recommended in case the final decision needs to be justified to either internal or external scrutiny in the future.

As with grouping, the criteria to assess whether the read-across is sufficiently justified depends on the purpose. The level of detail given in the worked examples in section 3.3.3, could be used as a guideline but the user should also include assessments of data quality (section 3.3.4) and similarity (section 3.3.5)

3.4.2 Worked Example: Use of read-across to satisfy a regulatory endpoint requirement for a group of MWCNTs

For this worked example we are referring to the MWCNT grouping exercise discussed in the "Using an IATA" section (3.3.3.2).

1. **Review purpose and data matrix to identify data gaps that need to be filled by read-across**

The purpose of grouping was regulatory and was specifically aimed at satisfying the REACH endpoint for chronic inhalation toxicity for a number of different MWCNTs.

2. **Generate a read-across hypothesis**

The grouping hypothesis that was tested was a GRACIOUS pre-defined hypothesis H-I-1, which states "*Respirable, bio persistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of NFs can occur resulting in lung toxicity*".

Assessment of the data matrix shows that OECD TG 413, which will satisfy the requirements of repeated dose exposure via inhalation for a registration under REACH, is available for MWCNT-E.

The read-across hypothesis can be written as:

The requirement of the repeated dose toxicity via inhalation for REACH registration will be satisfied for all group members by reading across OECD TG 413 data from MWCNT-E to all other NFs of MWCNT in the group. This is justified because a similarity assessment has confirmed that all group members are respirable, bio persistent, rigid HARNs so, following inhalation exposure, long-term pulmonary retention of NFs can occur resulting in lung toxicity.

3. **Substantiation – Check proposed source NF or non-NF is part of the group and generate further data if required**

Following the use of IATA H-I-1 the following NFs of MWCNT could be grouped; MWCNT-B and MWCNT-E. The identified source NF, MWCNT-E is still part of the group so no further data needs to be generated for this hypothesis (note: a different grouping hypothesis was required for MWCNT-A, MWCNT-C and MWCNT-D)

4. **Assess read-across hypothesis**

As there are only two NFs in the group, the analogue approach (one-to-one method) will be used for the read-across.

5. **Fill data-gaps for target NFs**

There is not sufficient data to ascertain whether MWCNT-B or MWCNT-E is likely to display the most severe hazardous behavior from repeated exposure via inhalation as this would be largely judged from DN 6 and 7 (i.e. the form that shows the highest inflammatory response). Unfortunately there is currently no data from these studies. However, the current assessment is that the two forms will be sufficiently similar to allow the OECD TG 413 data for MWCNT-E to be applied directly to MWCNT-B. In a one-to-one read-across approach this can be described as both a "*most conservative value"* method or a "*copying from one source NF*".

6. **Document read-across justification**

The details given in Section 3.3.3.2 would be a good basis for documenting the grouping and read-across justification.

3.4.3 Do you want to know more?

The following resources can provide more information:

ECHA, (2008). Guidance on information requirements and chemical safety assessment Chapter R.6: QSARs and grouping of chemicals. Available at https://echa.europa.eu/documents/10162/17224/information_requirements_r6_en.pdf/77f49f81-b76d-40ab-8513-4f3a533b6ac9

Guidance on application of read-across to all substance for registration under REACH.

ECHA (2012). Guidance on information requirements and chemical safety assessment. Chapter R.8: Characterisation of dose [concentration]-response for human health

https://echa.europa.eu/documents/10162/17224/information_requirements_r8_en.pdf/e153243a-03f0-44c5-8808-88af66223258

Section R.8.7 explains how a user can assess multiple pieces of data sources to identify the value that can be used as the source for the read-across.

ECHA (2017). Read-Across Assessment Framework (RAAF). Available at https://echa.europa.eu/documents/10162/13628/raaf_en.pdf/614e5d61-891d-4154-8a47-87efebd1851a

The regulation that specifies the methods to be used for read-across according to REACH.

ECHA (2019). Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals. In: Guidance on information requirements and chemical safety assessment. European Chemicals Agency.

Available at https://echa.europa.eu/documents/10162/2324906/appendix_r6_nanomaterials_en.pdf/71ad76f0-ab4c-fb04-acba-074cf045eaaa

The guidance provided by ECHA to apply the Read-Across Assessment Framework to nanomaterials.

Stone, V., Gottardo, S., Bleeker, E., Braakhuis, H., Dekkers, S., Fernandes, T., Haase, A., Hunt, N., Hristozov, D., Jantunen, P., Jeliazkova, N., Johnston, H., Lamon, L., Murphy, F., Rasmussen, K., Rauscher, H., Jiménez, A. S., Svendsen, C., Spurgeon, D., Oomen, A. G. (2020). A framework for grouping and read-across of nanomaterials- supporting innovation and risk assessment. Nano Today, 35, [100941]. doi.org/10.1016/j.nantod.2020.100941

This paper gives an introduction to the GRACIOUS Framework

3.5 Introduction to the tools to assist with the Framework

3.5.1 GRACIOUS Blueprint

The blueprint is a document intended for software developers who want to implement the GRACIOUS Framework, or parts of it, into their software product. Other users of the Framework are more likely to be exposed to the Blueprint as they use these software products. Therefore, this section is of interest mainly to software developers.

As the GRACIOUS Framework is highly complementary to risk assessment frameworks, integration of it into existing risk assessment tools or future SbD tools is considered the preferred way of automation. The highly user-interactive nature of the GRACIOUS Framework is less suited to be provided as a direct software service to be used via an API. Therefore, integration of relevant GRACIOUS Framework parts into external tools by programming them directly into the source code is more obvious. However, these software tools differ in their scope, scale, covered functionalities, modelled rules, used properties, terminology and definitions, user-interfaces and last but not least the programming languages and implementation techniques used to develop them.

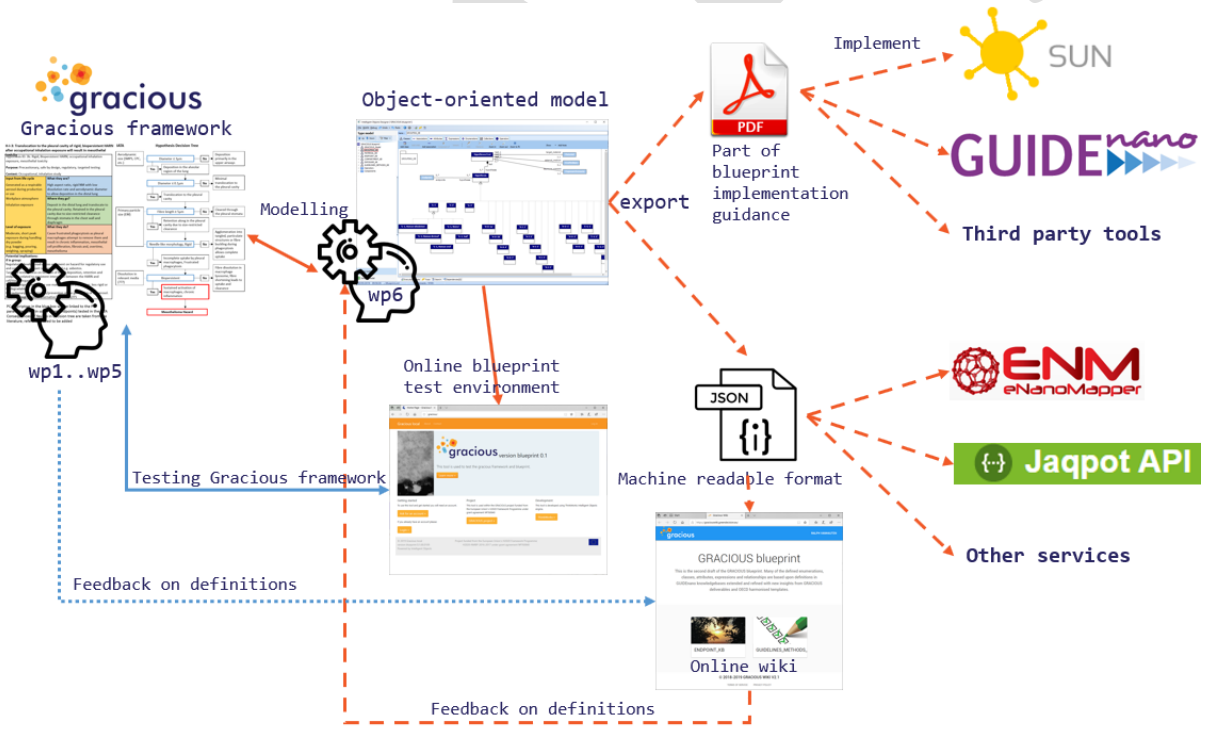
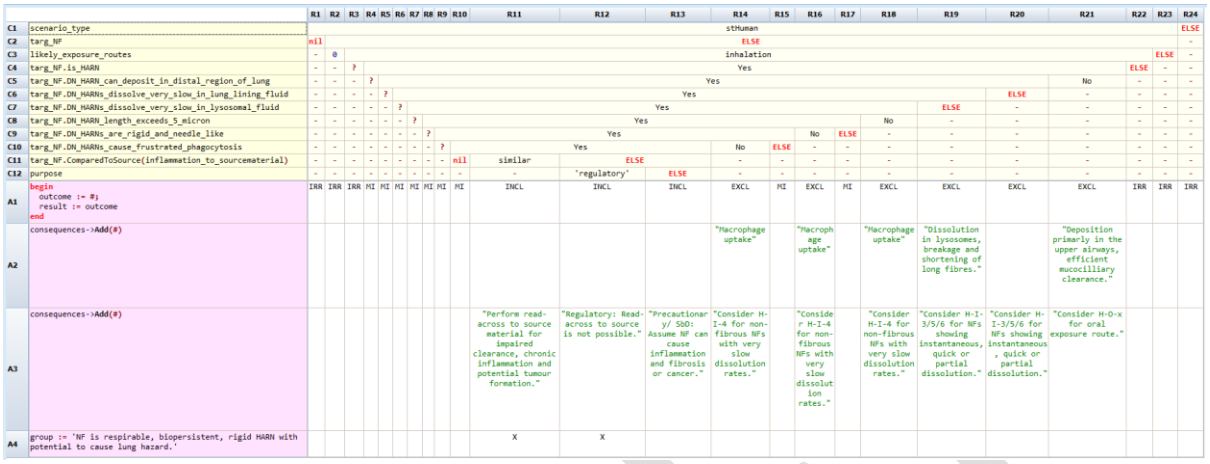


Figure 3.5.1: Information flow diagram showing the iterative development of the GRACIOUS Blueprint and how it allows integration of the GRACIOUS Framework with other tools

In order to help software developers with this integration process, the GRACIOUS Blueprint has been developed and will become available to developers at the end of the project (Figure 3.5.1). The Blueprint

describes the GRACIOUS Framework as an object-oriented model and contains most of the generic decision logic used in the Framework described in decision tables and algorithms in a pseudo code manner. The Blueprint document is automatically generated from the object-oriented model and as such a direct representation of it (Figure 3.5.2).



		R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	R12	R13	R14	R15	R16	R17	R18	R19	R20	R21	R22	R23	R24
C1	scenario_type	stHuman																							ELSE
C2	targ_NF	nil	ELSE																						-
C3	likely_exposure_routes	-	@	inhalation																				ELSE	-
C4	targ_NF.is_HARN	-	-	?	Yes																		ELSE	-	-
C5	targ_NF.DN_HARN_can_deposit_in_distal_region_of_lung	-	-	-	?	Yes																No	-	-	-
C6	targ_NF.DN_HARNs_dissolve_very_slow_in_lung_lining_fluid	-	-	-	-	?	Yes														ELSE	-	-	-	-
C7	targ_NF.DN_HARNs_dissolve_very_slow_in_lysosomal_fluid	-	-	-	-	-	?	Yes												ELSE	-	-	-	-	-
C8	targ_NF.DN_HARN_length_exceeds_5_micron	-	-	-	-	-	-	?	Yes										No	-	-	-	-	-	-
C9	targ_NF.DN_HARNs_are_rigid_and_needle_like	-	-	-	-	-	-	-	?	Yes							No	ELSE	-	-	-	-	-	-	-
C10	targ_NF.DN_HARNs_cause_frustrated_phagocytosis	-	-	-	-	-	-	-	-	?	Yes				No	ELSE	-	-	-	-	-	-	-	-	-
C11	targ_NF.ComparedToSource(inflammation_to_sourcematerial)	-	-	-	-	-	-	-	-	-	nil	similar	ELSE		-	-	-	-	-	-	-	-	-	-	-
C12	purpose	-	-	-	-	-	-	-	-	-	-	-	'regulatory'	ELSE	-	-	-	-	-	-	-	-	-	-	-
A1	begin outcome := #; result := outcome end	IRR	IRR	IRR	MI	MI	MI	MI	MI	MI	MI	INCL	INCL	INCL	EXCL	MI	EXCL	MI	EXCL	EXCL	EXCL	EXCL	IRR	IRR	IRR
A2	consequences->Add(#)														"Macrophage uptake"		"Macrophage uptake"		"Macrophage uptake"	"Dissolution in lysosomes, breakage and shortening of long fibres."		"Deposition primarily in the upper airways, efficient mucociliary clearance."			
A3	consequences->Add(#)											"Perform read-across to source material for impaired clearance, chronic inflammation and potential tumour formation."	"Regulatory: Read-across to source is not possible."	"Precautionary/ SbD: Assume NF can cause inflammation and fibrosis or cancer."	"Consider H-I-4 for non-fibrous NFs with very slow dissolution rates."		"Consider H-I-4 for non-fibrous NFs with very slow dissolution rates."		"Consider H-I-4 for non-fibrous NFs with very slow dissolution rates."	"Consider H-I-3/5/6 for NFs showing instantaneous, quick or partial dissolution."	"Consider H-I-3/5/6 for NFs showing instantaneous, quick or partial dissolution."	"Consider H-O-x for oral exposure route."			
A4	group := 'NF is respirable, biopersistent, rigid HARN with potential to cause lung hazard.'											x	x												

Figure 3.5.2: Example of how the Inhalation IATA, H-I-1, is modelled and visualised within the GRACIOUS Blueprint.

As development and revisions of the GRACIOUS Framework and content such as the IATAs and read-across and similarity approaches were ongoing almost until the end of the project the Blueprint does not cover all of the Framework content in detail. Nevertheless, the Blueprint will significantly assist software developers in identifying and understanding the required structures and rules to be implemented in their software. Publicly available GRACIOUS deliverables and publications can further assist them in incorporating the details. Integration of aspects of the Blueprint is underway or planned in both SUNDS and GUIDEnano via the projects SUNSHINE and SyByNa (see section 3.5.3). The Blueprint also contains an extensive network of descriptors for endpoints, assays, media, contributing activities etc. which can be used to correctly map software specific properties onto the ones used in the GRACIOUS framework (see GRACIOUS wiki).

Blueprint test-environment

Modelling of the Blueprint started in parallel to the development of the GRACIOUS Framework and also contributed to the alignment of the IATAs, DNs and descriptors used in them. The modelling was done in an object-oriented knowledge modelling environment. To ensure that the Blueprint model reflects the workflow of the GRACIOUS Framework as described in this guidance document, and to verify that the Blueprint knowledge models can actually form the basis for executable software components, an online Blueprint test-environment was developed. This test-environment is a user-interface directly interacting with the Blueprint object-model and decision logic. It allows testing the Blueprint by building up a case and indicating its purpose and providing Basic Information for the candidate NFs and potential use scenarios. Based on the provided information, the relevant IATAs are triggered and the outcome of the DNs within them can be determined in a tiered manner (Figure 3.5.3). Finally, the test-environment generates IATA-based data matrices as input for the similarity and grouping algorithms.

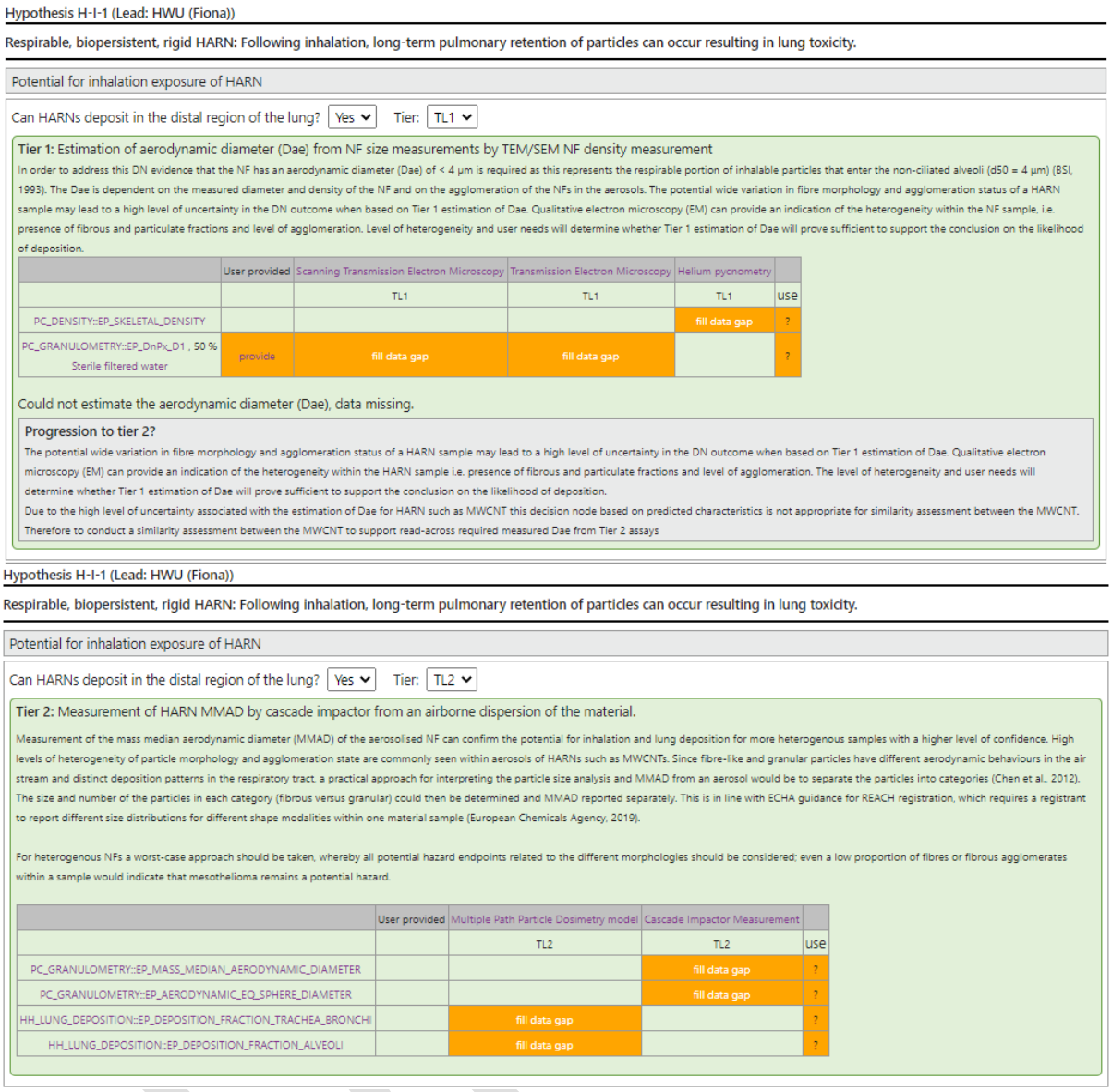
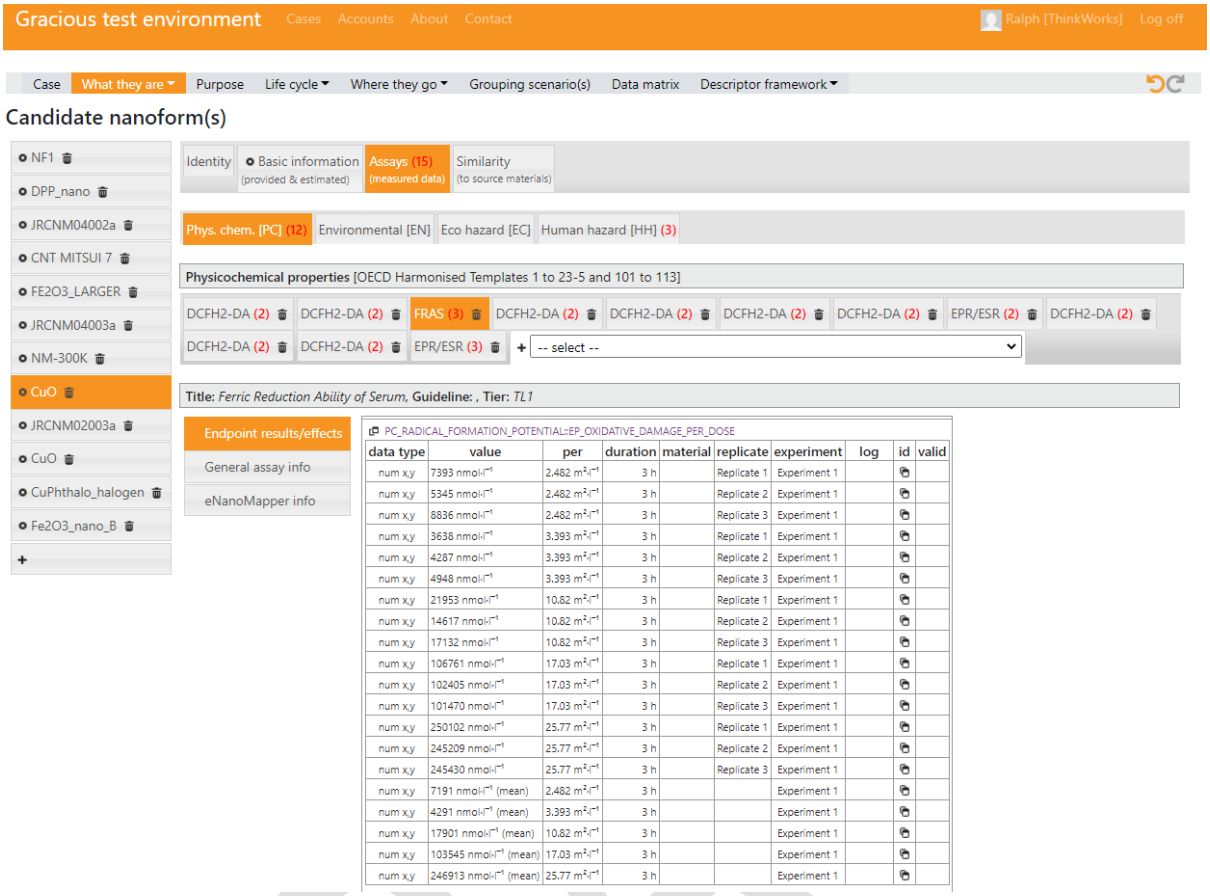


Figure 3.5.3: Screenshot of the GRACIOUS Blueprint test environment showing tier 1 and tier 2 of Decision Node 1 of the inhalation IATA, H-I-1

The test-environment was used to check and support the Blueprint development in an iterative way. The test-environment was also used to test the interoperability of the GRACIOUS Framework with the GRACIOUS database hosted by an eNanoMapper instance (Figure 3.5.4).



data type	value	per	duration	material	replicate	experiment	log	id	valid
num x,y	7393 nmol·l^{-1}	2.482 m^2·l^{-1}	3 h		Replicate 1	Experiment 1			
num x,y	5345 nmol·l^{-1}	2.482 m^2·l^{-1}	3 h		Replicate 2	Experiment 1			
num x,y	8836 nmol·l^{-1}	2.482 m^2·l^{-1}	3 h		Replicate 3	Experiment 1			
num x,y	3638 nmol·l^{-1}	3.393 m^2·l^{-1}	3 h		Replicate 1	Experiment 1			
num x,y	4287 nmol·l^{-1}	3.393 m^2·l^{-1}	3 h		Replicate 2	Experiment 1			
num x,y	4948 nmol·l^{-1}	3.393 m^2·l^{-1}	3 h		Replicate 3	Experiment 1			
num x,y	21953 nmol·l^{-1}	10.82 m^2·l^{-1}	3 h		Replicate 1	Experiment 1			
num x,y	14617 nmol·l^{-1}	10.82 m^2·l^{-1}	3 h		Replicate 2	Experiment 1			
num x,y	17132 nmol·l^{-1}	10.82 m^2·l^{-1}	3 h		Replicate 3	Experiment 1			
num x,y	106761 nmol·l^{-1}	17.03 m^2·l^{-1}	3 h		Replicate 1	Experiment 1			
num x,y	102405 nmol·l^{-1}	17.03 m^2·l^{-1}	3 h		Replicate 2	Experiment 1			
num x,y	101470 nmol·l^{-1}	17.03 m^2·l^{-1}	3 h		Replicate 3	Experiment 1			
num x,y	250102 nmol·l^{-1}	25.77 m^2·l^{-1}	3 h		Replicate 1	Experiment 1			
num x,y	245209 nmol·l^{-1}	25.77 m^2·l^{-1}	3 h		Replicate 2	Experiment 1			
num x,y	245430 nmol·l^{-1}	25.77 m^2·l^{-1}	3 h		Replicate 3	Experiment 1			
num x,y	7191 nmol·l^{-1} (mean)	2.482 m^2·l^{-1}	3 h			Experiment 1			
num x,y	4291 nmol·l^{-1} (mean)	3.393 m^2·l^{-1}	3 h			Experiment 1			
num x,y	17901 nmol·l^{-1} (mean)	10.82 m^2·l^{-1}	3 h			Experiment 1			
num x,y	103545 nmol·l^{-1} (mean)	17.03 m^2·l^{-1}	3 h			Experiment 1			
num x,y	246913 nmol·l^{-1} (mean)	25.77 m^2·l^{-1}	3 h			Experiment 1			

Figure 3.5.4: Screenshot of GRACIOUS Blueprint test environment showing how data can be imported from databases. FRAS assay data for CuO being loaded from eNanomapper and mapped onto the GRACIOUS descriptors

The test environment also contains functionality that allows visualisation of data in graphical form (Figure 3.5.5).

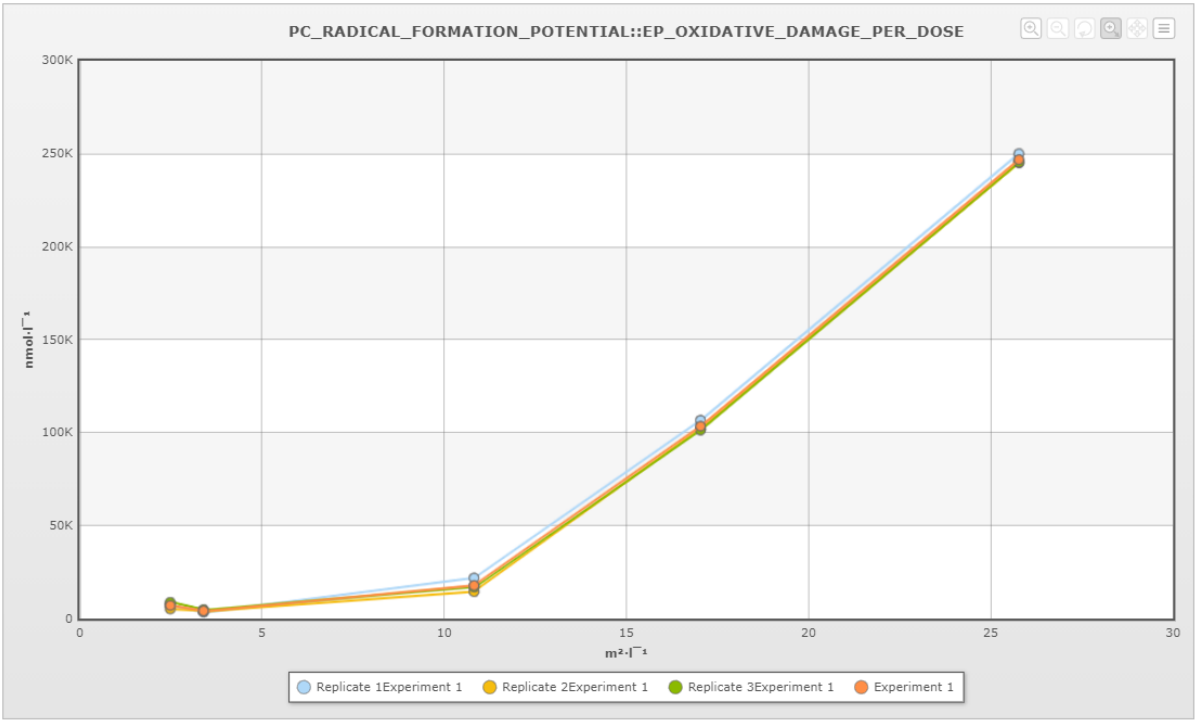


Figure 3.5.5: A screenshot from the GRACIOUS Blueprint test environment showing the response curve of dose versus radical formation potential showing how it can be used to visualise experimental results.

Any user of the (future) tools that integrate the Blueprint must be aware that they need to apply their own experience to the interpretation of results and drawing conclusions on Grouping and how to proceed once Grouping is justified.

3.5.2 GRACIOUS Wiki

The GRACIOUS Wiki was established to ensure that terminology was used in a consistent fashion across all parties involved in the development of the GRACIOUS Framework. It is now hosted within the Terminology Harmonizer developed by GreenDecisions (https://terminology-harmonizer.greendecision.eu/), where it is joined by similar developments from other projects. The Wiki is divided into 6 sections that cover all aspects of the Framework (Figure 3.5.6).

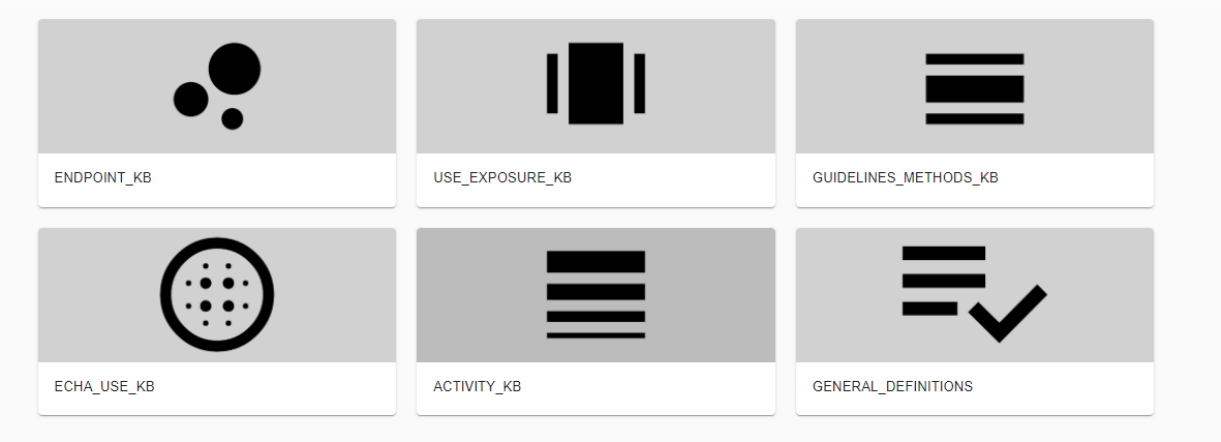


Figure 3.5.6: Screenshot of the frontpage of the GRACIOUS Wiki showing the sections of the GRACIOUS Framework addressed.

Within these sections a user can identify a specific descriptor for endpoints, assays, media, cell lines, etc. used in the GRACIOUS Framework, see its definition and where the term and definition have been used in other circumstances (Figure 3.5.7).

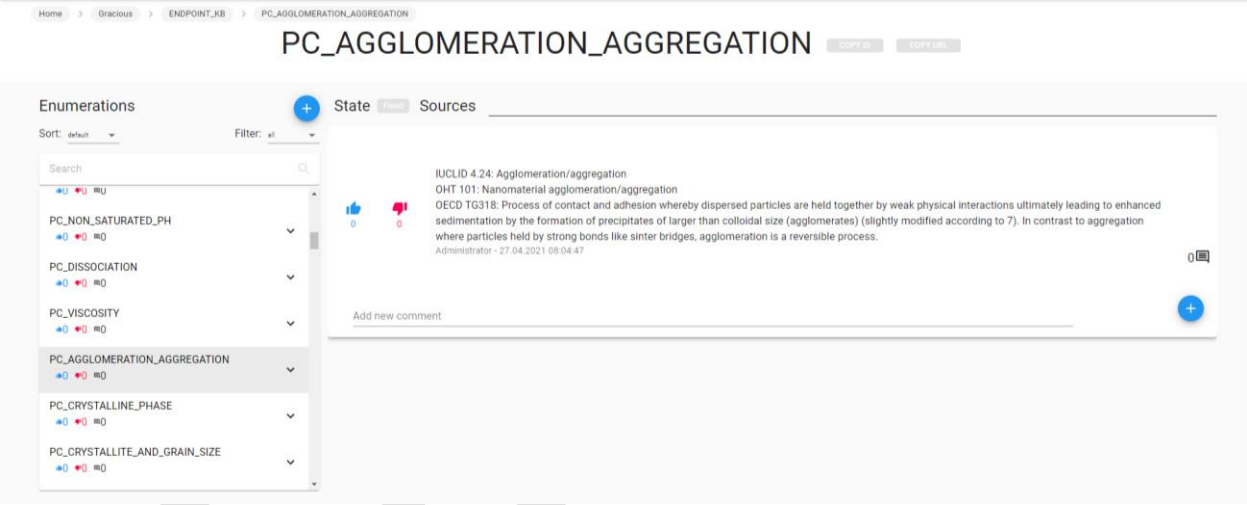


Figure 3.5.7: Screen shot of the GRACIOUS Wiki giving the definition of the term Agglomeration/Aggregation and from where this definition has been derived.

Users are able to make comments on the terms and definitions and to suggest new ones to be added.

3.5.3 Do you want to know more?

On request, the online test-environment can be made accessible to software developers to illustrate how the different elements of the Gracious Framework could be presented in a user-interface and to support integration efforts.

The Blueprint document will be made available as deliverable D6.7 from the GRACIOUS project.

Guidenano (https://www.guidenano.eu/)

GUIDEnano is a European research project funded under the 7th framework programme developing a web-based guidance tool, which will help the nano-enabled products users to design and apply the most appropriate risk assessment & mitigation strategy for a specific product.

SAbyNA (https://www.sabyna.eu/)

SAbyNA aims to improve the usability of existing databases, test methods, models, frameworks and tools and integrate them into an interactive and user-friendly web-based guidance.

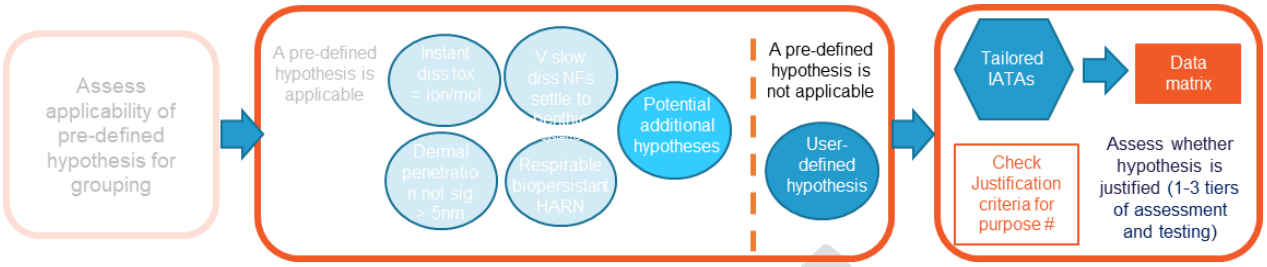
SUNDS (https://sunds.gd/)

Decision support system for risk assessment and management of nano(bio)materials used in consumer products and medical applications

SUNSHINE (https://www.h2020sunshine.eu/)

SUNSHINE is an industry-oriented project, where leading research and technology organisations will cooperate with SMEs and large industries to develop and implement simple, robust, and cost-effective Safe and Sustainable by Design (SSbD) strategies for materials and products incorporating advanced multi-component nanomaterials. To this end, the project will establish a user-friendly e-infrastructure to foster dialogue, collaboration, and information exchange between actors along entire product supply chains.

3.6 Writing a user-defined hypotheses



It is felt that the pre-defined hypotheses developed by the GRACIOUS project should cover most issues that a user could encounter in the current scientific and regulatory environment. However, it is possible that new toxicological concerns will arise in the future, so this section is intended to highlight the key principles that should be kept in mind if a user needs to write their own hypothesis and associated IATA.

3.6.1 Instructions for writing a new hypothesis and IATA

The basic steps that need to be addressed when writing a user-defined hypothesis and constructing its associated IATA are shown below. A user should be aware that if they are writing their own hypothesis, it will be necessarily unique to them, so they will need to adapt the details of each step to their own situation.

Hypothesis

1. Identify the hazard of concern and the route of exposure that could lead to the adverse effect.
2. Understand the Mechanism/Mode of Action (MoA) linked to the hazard – **"What they do"**
3. Recognise how the NF is taken up by, distributed within and excreted from the target organism or its immediate environment – **"Where they go"**
4. Identify the physicochemical characteristics linked with the MoA – **"What they are"**

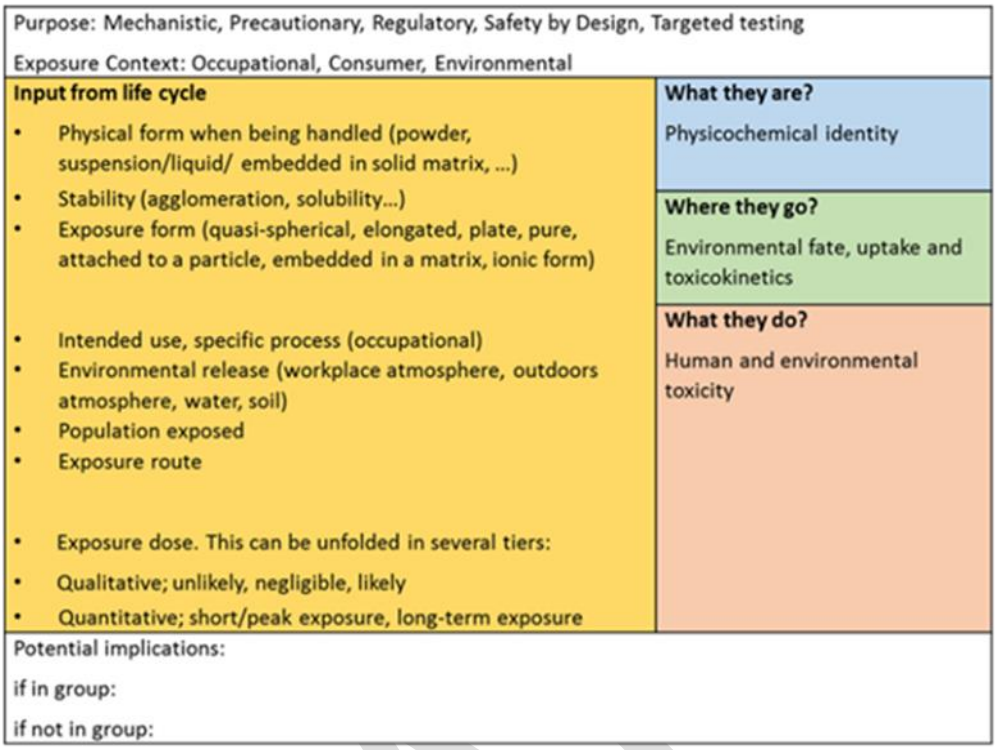


Figure 3.6.1: Template of the matrix used by the GRACIOUS project to compile relevant information when writing a grouping hypothesis.

The GRACIOUS project found using the template in Figure 3.6.1 useful to collate all the relevant information to design a grouping hypothesis in one place.

IATA

1. Identify DNs that investigate critical parameters of the hypothesis, allowing candidate NFs that do not meet the criteria to be removed from the proposed group. DNs are worded as questions about the candidate nanoforms. The DN should address each of the key aspects of the hypothesis and are usually arranged within the IATA in the following order.
    a. What they are
    b. Where they go
    c. What they do
2. Identify studies that can detect or measure the critical parameters associated with each DN.
3. Assign each study to a tier within the DN based on its complexity, resource intensiveness or whether it requires *in vivo* studies.
    a. Tier 1 studies are often physicochemical, *in silico,* simple *in vitro* or acute invertebrate studies.
    b. Tier 2 studies can be more complex physicochemical or *in vitro* studies, longer-term invertebrate studies or very simple *in vivo* studies.
    c. Tier 3 studies are usually *in vivo* studies for human health, or mesocosym studies for environmental toxicity.

It is not obligatory to identify studies for every tier within a DN.

4. Decide whether a threshold or floating band defined by similarity will be used to reach a conclusion for each tier and for each DN.
5. Examine which similarity methodology will be used to draw a final conclusion.

We hope that the detailed explanations of the pre-defined hypotheses in previous sections will give the user a background into the thought processes that were used to construct them.

Tips for new users

Most IATAs can be made by editing an existing IATA. For example, if the route of exposure is inhalation, look at all of the inhalation IATAs and prioritise the IATA which includes DNs relevant to the hypothesis to be tested. Delete or edit the DNs that are not relevant and add new DNs where required.

4.0 Annexes

4.1 Environmental Hypotheses

Identifier	Description
E-G-1a	NFs with a quick dissolution rate in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the solutes.
E-G-1b	NFs with a very slow dissolution rate and a stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the NFs in aqueous environment.
E-G-1c	NFs with a very slow dissolution rate and a partial stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of the particles in aqueous environments.
E-G-1d	NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (a high toxicity ratio solute : NF allows read-across to similar solutes).

E-G-1e	NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (a low toxicity ratio solute : NF allows read-across to similar NFs).
E-G-1f	NFs that partially dissolve in a (partial) stable dispersion in environmentally relevant aquatic media: Following aqueous exposure lethal and sub-lethal toxicity to representative aquatic species is driven by the fate and toxicity characteristics of both NF particles and solutes in aqueous environments (an intermediate toxicity ratio solute : NF limits possibilities for read-across).
E-G-2	NFs with a very slow dissolution rate in environmentally relevant media: Bio persistence potential is likely which triggers concern for long-term lethal and sub-lethal toxicity to representative environmental species.
E-G-3a	NF for which dissolution products are chemically transformed into a “new” NF as a result of speciation with the surrounding medium: hazards are driven by the fate and hazard characteristics of the “new” bio persistent NF
E-G-3b	NF that are chemically transformed into a “new” persistent NF as a result of speciation with the surrounding medium: hazards are driven by the hazard characteristics of the “new” NF.
E-G-4a	NFs with an organic surface treatment that is lost from the NF surface following exposure in WWTP compartment can be grouped: Fate and toxicity of the exposure relevant NF can be considered similar to a non-coated analogous NF in WWTP compartment.

E-G-4b	NFs with an organic surface treatment that is lost from the NF surface following exposure in aquatic compartment can be grouped: Fate and toxicity of the exposure relevant NF can be considered similar to a non-coated analogous NF in aquatic compartment.
E-G-4c	NFs with an organic surface treatment that is lost from the NF surface following exposure in soil compartment can be grouped: Fate and toxicity of the exposure relevant NF can be considered similar to a non-coated analogous NF in soil compartment
E-G-4d	NFs with a durable and toxic organic surface treatment cannot be grouped for read-across in any environmental/system compartment: Specific testing is required.
E-WS-1a	Bioavailable NFs with a very slow dissolution rate in sediment can be grouped: Following sediment exposure, NFs in this group will maintain nano-specific activity and can cause lethal and sub-lethal toxicity to representative benthic species.
E-WS-1b	Bioavailable NFs with a quick dissolution rate in sediment can be grouped: Following sediment exposure, the dissolution products of NFs in this group can cause lethal and sub-lethal toxicity to representative benthic species.
E-WS-1c	Bioavailable NFs that partially dissolve in sediment and have a low toxicity ratio dissolution products: NF can be grouped: Following sediment exposure, NFs in this group can cause lethal and sub-lethal toxicity to representative benthic species.
E-WS-1d	Bioavailable NFs that partially dissolve in sediment and have a high toxicity ratio dissolution products: NF can be grouped: Following sediment exposure, the dissolution products of NFs in this group can cause lethal and sub-lethal toxicity to representative benthic species.
E-WS-1e	Bioavailable NFs that partially dissolve in sediment and have an intermediate toxicity ratio dissolution products: NF can be grouped: Following sediment exposure, NFs in this group,

	together with their dissolution products can cause lethal and sub-lethal toxicity to representative benthic species.
E-S-1a	NFs with a very slow dissolution rate and low affinity with the solid soil phase: Following soil exposure NF mobility in soil follows ground water flows. NFs in this group can cause acute lethal and sub-lethal toxicity to representative soil species.
E-S-1b	NFs with a very slow dissolution rate and high affinity with the solid soil phase: Following soil exposure persistence in soil is likely. NFs in this group can cause (long-term) lethal and sub-lethal toxicity to representative soil species.
E-S-2a	NFs with a quick dissolution rate in soil: Following soil exposure lethal and sub-lethal toxicity to representative soil species is driven by the fate and toxicity characteristics of the dissolution products.
E-S-2b	NFs that partially dissolve in soil: Following soil exposure fate is driven by dissolution rather than mobility of the NFs. Hazard will be driven by the contribution of solutes and particles to the overall toxicity of the exposure

4.2 Human Health Hypotheses

Identified	Description
H-I-1	Respirable, bio persistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of HARNs can occur resulting in lung toxicity.
H-I-2	Respirable, bio persistent, rigid HARN: Following inhalation exposure and translocation of HARNs to the pleura, mesothelioma development can occur.
H-I-I	Respirable NFs showing instantaneous dissolution: Following inhalation exposure, the toxicity is driven by and is therefore similar to those of the constituent ions or molecules.

H-I-Q	Respirable NFs showing quick dissolution: Following inhalation both NFs and constituent ions or molecules may contribute to toxicity, but there is no concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release.
H-I-G	Respirable NFs showing gradual dissolution: Following inhalation exposure both NFs and constituent ions or molecules may contribute to toxicity and there is some concern for accumulation. Toxicity (also) depends on the location of the ionic or molecular release.
H-I-S	Respirable NFs showing very slow dissolution: Following inhalation exposure, toxicity is driven by the NFs and accumulation of NFs in the lungs can lead to long-term toxicity.
H-O-I	NFs with an instantaneous dissolution : Following oral exposure, the toxicity is driven by and is therefore similar to that of the constituent ions or molecules.
H-O-Q1	NFs with a quick dissolution: Following oral exposure both NFs and constituent ions or molecules may contribute to local inflammation in the OGI tract, but there is no concern for NF accumulation
H-O-Q3	NFs with a quick dissolution: Following oral exposure both NFs and constituent ions or molecules may drive antimicrobial impacts (e.g., reducing microbial content and diversity within the OGI tract), but there is no concern for NF accumulation.
H-O-G1	NFs showing gradual dissolution: Following oral exposure both NFs and constituent ions or molecules may lead to local inflammation in the GIT.
H-O-G2	NFs showing gradual dissolution : Following oral exposure both NFs and constituent ions or molecules may translocate to secondary target organs and may lead to systemic toxicity in secondary organs.

H-O-G2	NFs showing gradual dissolution : Following oral exposure both NFs and constituent ions or molecules may drive antimicrobial impacts, such as reducing microbial content and diversity within the GIT.
H-O-S1	NFs with a very slow dissolution rate: Following oral exposure NFs will maintain nanospecific activity that may lead to local inflammation within the GIT.
H-O-S2	NFs with a very slow dissolution rate: Following oral exposure NFs will maintain nanospecific activity that may drive translocation across the GIT wall, subsequent bio persistence in the body and systemic toxicity in secondary organs.
H-O-S3	NFs with a very slow dissolution rate : Following oral exposure NFs will maintain nanospecific activity that will drive antimicrobial impacts, such as reducing microbial content and diversity within the GIT
H-D-1	NFs with an instantaneous dissolution: Following dermal exposure NFs will dissolve into their molecular or ionic form before they reach the viable layers of the skin and will cause similar toxicity as substances quickly releasing, dissolving and/or transforming into the same ionic or molecular forms.
H-D-2	NFs with constituent substance(s) or degradation products classified for dermal irritation or sensitization: Dermal exposure to the NFs may result in dermal irritation or sensitization.
H-D-3	NFs that are not bio persistent : Dermal exposure to NFs will not lead to accumulation of NFs or subsequent systemic toxicity.
H-D-4	NFs that are not flexible and have a constituent particle size larger than 5 nm: Following dermal exposure NFs will result in limited or no dermal absorption.