ALTEX, accepted manuscript
published June 22, 2022, version 3
doi:10.14573/altex.2203241

Concept Article

How to Formulate Hypotheses and IATA to Support Grouping and Read-Across of Nanoforms

Fiona A. Murphy[1], Helinor J. Johnston[1], Susan Dekkers[2], Eric A. J. Bleeker[2], Agnes G. Oomen[2,3], Teresa F. Fernandes[1], Kirsten Rasmussen[4], Paula Jantunen[4], Hubert Rauscher[4], Neil Hunt[5], Luisana di Cristo[6], Hedwig M. Braakhuis[2], Andrea Haase[7], Danail Hristozov[8], Wendel Wohlleben[9], Stefania Sabella[6] and Vicki Stone[1]

[1]Institute of Biological Chemistry, Biophysics and Bioengineering, Heriot-Watt University, Edinburgh, UK; [2]National Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands; [3]Institute for Biodiversity & Ecosystem Dynamics (IBED), Faculty of Science, University of Amsterdam, Amsterdam, The Netherlands; [4]European Commission, Joint Research Centre (JRC), Ispra, Italy; [5]Yordas Group, Lancaster Environment Centre, Lancaster University, Lancaster, UK; [6]Nanoregulatory Group, Drug Discovery and Development Department, Italian Institute of Technology, Genova, Italy; [7]German Federal Institute for Risk Assessment (BfR), Department of Chemicals and Product Safety, Berlin, Germany; [8]GreenDecision, Venice, Italy; [9]BASF SE, Dept. Material Physics and Dept of Experimental Toxicology & Ecology, Ludwigshafen, Germany

Abstract

Manufacturing and functionalizing materials at the nanoscale has led to the generation of a whole array of nanoforms (NFs) of substances varying in size, morphology and surface characteristics. Due to financial, time and ethical considerations, testing every unique NF for adverse effects is virtually impossible. Use of hypothesis-driven grouping and read-across approaches, as supported by the GRACIOUS Framework, represent a promising alternative to case-by-case testing which will make the risk assessment process more efficient. Through application of appropriate grouping hypotheses, the Framework facilitates the assessment of similarity between NFs, thereby supporting grouping and read-across of information, minimizing the need for new testing and aligning with 3R principles of Replacement, Reduction and Refinement of animals in toxicology studies. For each grouping hypothesis an Integrated Approach to Testing and Assessment (IATA) guides the user in data gathering and acquisition to test the hypothesis, following a structured format to facilitate efficient decision-making. Here we present the template used to generate the GRACIOUS grouping hypotheses encompassing information relevant to "Lifecycle, environmental release and human exposure", "What they are: physicochemical characteristics", "Where they go: Environmental fate, uptake and toxicokinetics", and "What they do: human and environmental toxicity". A summary of the template-derived hypotheses focusing on human health is provided, along with an overview of the IATAs generated by the GRACIOUS project. We discuss the application and flexibility of the template, providing the opportunity to expand the application of grouping and read-across, in a logical evidence-based manner, to a wider range of NFs and substances.

1 Introduction

Nanotechnology is a key enabling technology of the 21st century, with the exploitation of nanomaterials (NMs) in a range of applications associated with many benefits for society and the global economy[1]. However, uncertainties regarding the possible risks posed by NMs to human health and the environment need to be better understood in order to expedite the success of this technology. Identification of the possible adverse effects posed by a broad range of NMs is critical to implement measures to manage any potential risks to human health and the environment (Aschberger et al., 2016; Bottero et al., 2015; Kuempel et al., 2012).

There can be many nanoforms (NFs) of the same substance that vary in, for example, size, morphology and surface characteristics. Under REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals) (EC, 2006; EC, 2018)

Disclaimer: The content expressed in this paper is solely the opinion of the authors and does not necessarily reflect the opinion of their institutions.

Received March 24, 2022; Accepted June 20, 2022;
Epub June 22, 2022;

ALTEX 39(#), ###-###. doi:10.14573/altex.2203241

Correspondence: Fiona Murphy, PhD
Heriot-Watt University
Riccarton Campus
The Avenue
Edinburgh EH14 4AS, UK
(f.murphy@hw.ac.uk)

[1] http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52012DC0341&locale=en

individual NFs of a nanomaterial are registered separately and as such safety assessment is required for each. Considering the NFs currently in production and under development, it is not feasible to perform an individual risk testing of every NF. In general, there is a global drive to move away from toxicological testing of substances, including NFs, using animal models (EC, 2010; Burden et al., 2017). This means that more effective approaches for risk assessment of NFs are needed, based on the 3Rs principles of Replacement, Reduction and Refinement[2]. Grouping and read-across approaches of chemicals is one option to fulfil regulatory information requirements (EC, 2006), and research also recommends such tools as useful for NF risk assessment (Lamon et al., 2019; Stone et al., 2020).

The concept of grouping is described in REACH as follows: "Substances whose physicochemical, toxicological and ecotoxicological properties are likely to be similar or follow a regular pattern as a result of structural similarity may be considered as a group, or 'category' of substances" (EC, 2006, Annex XI, 1.5). ECHA's Read-Across Assessment Framework guidance elaborates that establishment of a group allows the user to apply read-across. Read-across allows prediction of specific endpoint information for one or more substances (target material(s)) by using data for the same endpoint from another substance from within the same group (source material) for which more information is available (ECHA, 2017a). Using this approach, existing fate or hazard information on an appropriate source substance is used to predict the behavior of the new substance.

For chemicals, grouping and read-across are well-established concepts used to minimize the need for specific testing in risk assessment (OECD, 2017a, ECHA, 2008), and these approaches are already integrated into legislation, e.g., REACH (EC, 2006). With guidance provided by ECHA (ECHA, 2017b) and the OECD (OECD, 2016) on how to take into account nano-specific considerations, the application of grouping and read-across can be extended to the assessment of NFs. The amended REACH annexes state that 'for grouping different NFs of the same substance the molecular structural similarities alone cannot serve as a justification' (EC, 2018). In its nano-specific guidance, ECHA recommends applying the general similarity rules for grouping chemicals (Annex XI of REACH) to NFs. Furthermore, in addition to chemical composition and structure, more information about physiochemical (PC) parameters (e.g., aspect ratio, shape, solubility, surface area, surface charge, and surface treatment) are needed to support claims of similarity, or differences, between NFs (ECHA, 2017b). A review paper (Mech et al., 2019) gives an overview of the possibilities of applying grouping and read-across for the risk assessment of NFs in order to comply with the requirements of different pieces of EU legislation for various types of chemicals, highlighting the current challenges.

In order to streamline the implementation of grouping and read-across for NFs, the European Union's Horizon 2020 programme project GRACIOUS (full title: Grouping, Read-Across, CharacterIsation and classificatiOn framework for regUlatory risk assessment of manufactured nanomaterials and Safer design of nano-enabled products) developed a framework to facilitate the process (Stone et al., 2020).

The GRACIOUS Framework supports application of grouping and read-across for various purposes, defined by the user; for example, for filling data gaps when performing a risk assessment in a regulatory context (EC, 2006, Oomen et al., 2015). Additionally, grouping can facilitate the safe design of future NFs or nano-enabled products (NEPs), or lead to the adoption of precautionary measures to prevent human and environmental exposure.

Within the GRACIOUS Framework, each hypothesis will inform the content of a tailored Integrated Approach to Testing and Assessment (IATA). The IATAs guide the targeted gathering of existing and new information in a tiered manner, supporting the user in accepting or rejecting the grouping hypothesis.

This manuscript brings together the range of grouping hypotheses identified in GRACIOUS. In order to promote consistency in the approach to hypothesis generation, a grouping hypothesis template was developed in GRACIOUS and is provided here. It ensures the integration of PC properties, toxicokinetics and environmental fate in grouping for risk assessment. In addition, this manuscript brings together the different IATAs developed to provide the evidence needed to test each hypothesis and discusses how they have been applied. A strategy for the development of a tailored IATA as a direct extension from the hypothesis template is also discussed here so that the Framework can be further enriched over time for a wider array of NFs. Adoption and adaptation of the GRACIOUS hypothesis template is currently being explored by the next round of EU-funded H2020 projects (SUNSHINE[3], HARMLESS[4], DIAGONAL[5]) concerned with the development of strategies to facilitate the hazard assessment, grouping and read-across of more complex multicomponent and industrially relevant NMs and NEPs.

2 Structure of the grouping hypothesis template

The structure of the grouping hypothesis template is presented in Figure 1A. The template guides the user through a series of steps to develop a robust, evidence-based grouping hypothesis. The template outlines how to formulate a grouping hypothesis and indicates considerations that may influence the grouping decision. More specifically, we propose that the generation of a grouping hypothesis requires consideration of four main elements: 1) The rationale of the grouping (purpose and context), 2) Lifecycle/Exposure, 3) Interaction between PC characteristics, Fate and Hazard, and 4) Decision making (potential implications of the grouping).

[2] http://www.oecd.org/chemicalsafety/testing/animal-welfare.htm
[3] www.h2020sunshine.eu
[4] www.harmless-project.eu
[5] www.diagonalproject.eu

A

Hypothesis

Purpose: Mechanistic, Precautionary, Regulatory, Safety by Design, Targeted testing

Exposure Context: Occupational, Consumer, Environmental

Input from life cycle

Physical form when being handled (powder, suspension/liquid/ embedded in solid matrix, …)
Stability (agglomeration, solubility…)
Exposure form (quasi-spherical, elongated, plate, pure, attached to a particle, embedded in a matrix, ionic form)

Intended use, specific process (occupational)
Environmental release (workplace atmosphere, outdoors atmosphere, water, soil)
Population exposed
Exposure route

Level of Exposure

Exposure dose. This can be unfolded in several tiers:
Qualitative; unlikely, negligible, likely
Quantitative; short/peak exposure, long-term exposure

What they are

Physicochemical characteristics

Where they go

Environmental fate, uptake and toxicokinetics

What they do

Human and environmental toxicity

Potential implications:

if in group:

if not in group:

B

Hypothesis: Biopersistent HARN may pose an asbestos-like occupation inhalation hazard resulting in mesothelioma development

Purpose: Precautionary, Regulatory, Safety by Design, Targeted testing

Exposure Context: Occupational

Input from life cycle

Generated as a respirable aerosol during production or use
Workplace atmosphere
Inhalation exposure

Level of exposure

Moderate, short peak exposure during handling dry powder (e.g. bagging, pouring, weighing, spraying)

What they are

High aspect ratio, rigid NF with low dissolution rate and aerodynamic diameter to allow deposition in the distal lung

Where they go

Deposit in the distal lung and translocate to the pleural cavity. Retained in the pleural cavity due to size-restricted clearance through stomata in the chest wall and diaphragm

What they do

Cause frustrated phagocytosis as pleural macrophages attempt to remove them and result in chronic inflammation, mesothelial cell proliferation, fibrosis and, overtime, mesothelioma

Potential implications:

if in group:

Precautionary: limit exposure/prevention of the generation of an aerosol.

Regulatory: develop read-across argument on hazard for regulatory use and compare to relevant source material e.g. asbestos.

Safe-by-design: consider the use of materials which are shorter, less rigid or biodegradable.

Targeted testing: focus on location of HARN deposition, retention and endpoints related to persistent interaction between the HARN and cell/tissue/organism.

if not in group: consider alternative hypothesis

Biopersistent HARN that deposit in the distal lung and translocate to the pleural cavity may pose an asbestos-like occupational inhalation hazard resulting in mesothelioma development.

Fig. 1: The GRACIOUS grouping hypothesis template (A). An example of hypothesis generation (B)

2.1 Purpose and context

The initial step of generating a hypothesis for grouping with the help of the presented template is to define the 'purpose' of the grouping which may relate, for example, to regulatory, design and innovation, or precautionary needs of the user. This ensures the grouping will be relevant and of use to the user. Defining the purpose in the generation of the hypothesis will also impact future decisions made on the basis of the results from applying the grouping hypothesis, as discussed further below.

The 'context' identifies the use(s) of the NF, resulting in the identification of the likely exposure environment(s) (e.g., 'Occupational', 'Environment') and population(s) (e.g., 'Worker', 'Consumer') covered by the grouping hypothesis. This is the first step in defining the likely exposure scenarios (including the exposure route, duration, and frequency). It also determines whether one major route of exposure is of particular concern (e.g., inhalation, ingestion, release into the aquatic or terrestrial environment), or if multiple exposure scenarios need to be considered. The identification of multiple exposure scenarios will prompt the formulation of several separate context-dependent hypotheses, with different information needs, for the same NF. Initial consideration of the overall context will enable a focus to be placed on developing a grouping hypothesis for NF applications that may lead to high(er) human or environmental exposure and thereby, possibly, high(er) risks (Dekkers et al., 2016). Together, the purpose and context aspects of the template frame the rationale for the grouping hypothesis in a clear and specific manner.

2.2 Lifecycle and exposure scenario

During the lifecycle of a NF, many transformations may take place (which may alter some of the PC properties of the NF), and release may happen at several time-points. The human exposure scenario includes the route of exposure and the parameters that characterize specific population considerations (e.g., consumers, a vulnerable population, or workers), and the level,

frequency, and duration of the exposure. Considerations may also involve the release type (e.g., continuous or intermittent), and type of source (e.g., point source, multiple sources, diffuse releases from multiple (small) point sources). Exposure can be quantified by measurements or estimated by predictions. Taking lifecycle and exposure into account in grouping helps to identify NFs of concern in the most relevant real-life scenarios (Landsiedel, 2016) and is thus key to the development of a grouping hypothesis.

Within the template, the exposure scenario relevant for the defined purpose and context is based on input from lifecycle and release. This scenario informs the route of exposure, which leads to subsequent considerations of the PC characteristics, toxicokinetics and fate, and potential hazards relevant for grouping. Therefore, the inclusion of information on lifecycle and release when developing a grouping hypothesis will clarify the potential level and duration of exposure anticipated for the selected context either quantitatively or qualitatively (e.g., high, moderate or low), depending on the information available. If there is limited information available regarding the lifecycle or potential for release of a substance, the exposure scenario can be based on worst-case scenarios predicted for the processes under consideration, e.g., 100% of released NF will be inhaled.

The hypothesis template organises exposure data (predicted or measured) so that it links NF release to exposure at a target, which may be defined within the template at the level of a population or environmental compartment (soil, water, sediment, etc.), an individual organism, an organ or a cell. Knowledge of exposure when developing a grouping hypothesis is important, as it allows the identification of potential route(s) of exposure and subsequent target sites for NF accumulation, and links directly to the toxicokinetics and fate of the NF.

2.3 PC characteristics, fate, hazard - What they are, where they go, what they do

The sections ‘What they are’, ‘Where they go’ and ‘What they do’ (Stone et al., 2020) of the hypothesis template identify the considerations needed to establish a group (i.e. the linking of observed toxic responses to the PC properties of a NF, where possible). These sections concern the PC characterization, toxicokinetics and fate, and the (eco)toxicity of the NFs under investigation to assess similarity between proposed group members. The information included in these sections can also be used to define the boundaries of a group.

To support evidence-based grouping, the characterization of PC properties is required to both verify that the ‘target’ NF fits into the group related to the hypothesis (e.g. does the NF dissolve quickly in the relevant conditions?), and to allow comparison with other NFs in the group, including source materials (Arts et al., 2014). The ‘What they are’ section of the template specifically describes the PC characteristics shared between members of a group that determine the fate and hazard. For PC characterization of a NF, a distinction can be made between the intrinsic material properties of a NF (such as chemical composition, crystallinity, and water solubility) and system-dependent (extrinsic) properties defined by the surroundings in which the NF is placed (e.g. dissolution rate in biological media, surface reactivity and dispersibility) (Arts et al., 2014). Consideration of both intrinsic and extrinsic properties is important when developing a grouping hypothesis. For example, water solubility can be considered an intrinsic property as it is established in a specific medium under specific conditions. Solubility as such would not be considered an intrinsic property as it depends on the medium. Likewise, dissolution rate is highly dependent on the exact medium and medium conditions (temperature, whether the NF is residing in marine or fresh water, or in the lysosomal compartment of the cell etc.). These factors influence the distribution of the NF and potential for its accumulation in a target tissue or environmental compartment. To perform a relevant (eco)toxicological assessment, it is therefore important to consider both the intrinsic properties of the NF and the available knowledge regarding system-dependent properties when generating a grouping hypothesis (Sellers et al., 2015).

2.4 Potential implications

Considering the purpose of grouping when starting to design a grouping hypothesis allows the user to state explicitly the possible future actions based on the outcome of the hazard assessment, addressed in the ‘potential implications’ section of the template. Therefore, if the hypothesis is accepted, then action can be taken dependent on the reason why information on this endpoint is needed. For example, if the purpose of testing is to ensure the development of a safe(r) innovative NF or NEP, the potential implication of adhering to a hypothesis that causally links a specific PC characteristic to an adverse health effect may be to consider the impact of altering, if possible, a PC characteristic in the newly designed NF or NEP.

3 Using the template to generate a grouping hypothesis

An example of using the template to generate a grouping hypothesis is briefly described here; the details are fully elaborated in (Murphy et al., 2021). The specific example focuses on high aspect ratio nanoforms (HARNs), for which knowledge of fiber toxicology helps to inform the hazard endpoint and hypothesis.

Following inhalation, asbestos fibers can elicit adverse health impacts in the lung, including mesothelioma, a tumor that arises from the mesothelial cells lining the chest wall and lungs (Wagner et al., 1960). Mesothelioma is a disease almost exclusively associated with exposure to pathogenic fibers and rarely occurs as a spontaneous tumor, and it is thus considered a fiber-specific hazard (Selikoff et al., 1964). The asbestos literature has generated a structure/activity relationship (SAR) which defines the PC characteristics governing the pathogenicity of a fiber, known as the Fiber Pathogenicity Paradigm (FPP). Briefly, the FPP defines the intrinsic pathogenic properties of fibers, including a diameter which allows for deposition into the alveolar spaces and a length which prevents effective clearance from the lung and pleural cavity (Davis et al., 1986). The resistance of fibers to chemical dissolution or mechanical breakage within the biological milieu, so called biodurability, is an extrinsic factor dependent on the environment in which the fiber resides. Biodurability is also specified in the FPP as a critical determinant of fibre pathogenesis, as it leads to a build-up of dose in the tissue (Bernstein et al., 1994, 1997).

The FPP has been shown to be robust and hold true for a wide variety of fibers (Donaldson et al., 2010). HARNs (such as carbon nanotubes (CNTs)) can have the properties identified in the FPP, and have therefore been hypothesized to pose an asbestos-like inhalation hazard with the potential to cause mesothelioma in exposed populations (Nagai et al., 2011; Poland et al., 2008; Rittinghausen et al., 2014). Utilizing the hypothesis template, a HARN group can be constructed based on their shared potential to cause a common disease endpoint (mesothelioma) (Murphy et al., 2021). The grouping of potentially pathogenic HARNs may be useful for a number of purposes, including safe(r)-by-design innovation of new materials that do not lead to mesothelioma, supporting the adoption of precautionary hygiene measures and forming the basis of a read-across argument to enable waiving of tests for risk assessment. Based on the history of occupational inhalation exposure driving the epidemic of asbestos-related disease (Stayner et al., 2013), the chosen context for grouping HARNs based on their potential to cause asbestos-like hazard prioritizes the likelihood of inhalation exposure during the production of HARNs and HARN-containing materials. 'Lifecycle and exposure' will indicate the potential processes and scenarios resulting in release and aerosolization of HARNs and assessment of potential exposure levels. 'What they are' defines the relevant threshold dimensions (e.g., for length and diameter) for a HARN to meet the FPP criteria, determining where the HARN will deposit in the lung, and whether it can be effectively cleared from the lung and pleural cavity. Furthermore, 'What they are' and 'Where they go' together address the biopersistence of the HARNs and the likelihood of their significant retention in the target tissue of the distal lungs and mesothelium. 'What they do' frames the comparison between putative group members in the activation of relevant biological responses leading to fiber-related lung disease. The use of the template allows a structured hypothesis to be generated defining clear boundaries for the group and integrating the necessary considerations that may impact a grouping decision (Figure 1B).

The hypothesis is substantiated by extensive evidence from the asbestos literature (reviewed in (Broaddus et al., 2011; Craighead & Mossman, 1982; Donaldson et al., 2010). HARNs that meet the criteria set out by the boundaries of the group (threshold length, diameter, solubility level) can be considered members of this group. The carcinogenicity hazard (mesothelioma) posed by pathogenic fibers is caused by their structure, and therefore although the chemical composition can affect biopersistence, it is not specified as a critical factor for toxicity (Donaldson et al., 2010) or represented in the grouping hypothesis. This group could therefore contain structurally similar but chemically different substances, all of which exhibit biopersistence, ranging from different forms of naturally occurring mineral fibers to engineered HARNs. However, for regulatory purposes, for example under REACH, grouping based on the use of the HARN hypothesis would likely only be considered applicable for different NFs of the same chemical composition within a substance registration.

The potential implications of placing a NF in the HARN group are linked to the specific purpose of the grouping. For grouping purposes such as safe(r)-by-design, placing a HARN within this group may impact a cost/benefit analysis of whether the potential risks associated with the hazard endpoint specified by this group outweigh the potential economic or societal benefit of producing this HARN, or whether the design of the HARN should be amended to improve safety. For regulatory purposes, a direct comparison can be made between a data-rich source material from the group and the target in order to substantiate a read-across argument specific to the selected endpoint of carcinogenicity (mesothelioma). Rather than testing each new HARN using a standard 2-year carcinogenicity assay (OECD, 2018) for the potential to cause mesothelioma, comparison of the critical PC descriptors as defined by the hypothesis may be sufficient to build a read-across argument between the target NF and a known pathogenic HARN (source material) (Grosse et al., 2014). Well-substantiated read-across arguments resulting from an evidence-based grouping hypothesis may therefore be used in the future to support the regulatory classification of HARNs. For example, if a NF is currently considered not classifiable as to their carcinogenicity to humans (IARC Group 3) due to insufficient evidence, but it adheres to the HARN grouping hypothesis, then a decision could be made on its potential carcinogenicity which aligns with other members of the group for which sufficient evidence exists to support classification; and some CNTs are classified by IARC as possibly carcinogenic to humans (IARC Group 2B) (Cogliano et al., 2008; Grosse et al., 2014). Conversely, the same approach could be used to argue that a HARN does not adhere to the hypothesis defining mesothelioma as the endpoint, e.g. if the NF is not biopersistent, although grouping hypotheses for other hazard endpoints may still apply.

4 The GRACIOUS human health grouping hypotheses

The template has been used within the GRACIOUS project to develop a number of 'pre-defined' hypotheses for grouping and read-across (Tables 1-3). These grouping hypotheses were prioritized based on the amount of available information from literature available to substantiate the grouping hypothesis, as well as the clear implications of reaching a grouping decision.

Through application of the hypothesis template, six hypotheses relevant to exposure via inhalation (Table 1), nine hypotheses relevant to exposure via ingestion (Table 2) and four hypotheses relevant to dermal exposure (Table 3) were generated. Use of the template allows each hypothesis to identify initially the key physicochemical characteristics, followed by the relevant likely toxicokinetic pathway for the selected route of exposure, and then relates these to hazard endpoints.

For all three routes of exposure, dissolution rate of the NFs, and hence their biopersistence, is a key parameter identified in almost all hypotheses. This information is needed as it informs the users in the subsequent evidence gathering on whether they need to consider the solid form of the NF alone, the released ions/molecules or both in combination. The subsequent IATAs (described below) provide more details on the methods to be used to assess dissolution rate and the fluids in which dissolution assessment is relevant.

For the inhalation IATAs, the shape of particles is also key. The inhalation IATAs can be divided into those that address HARNs and follow the FPP (Murphy et al., 2021) as described above, and those that are not for HARNs (Braakhuis et al., 2021). Shape has not been identified as a key physicochemical characteristic for the oral or dermal routes of exposure. However, flexibility is indicated as a key PC characteristic in one of the dermal IATAs, since it relates to the ability of particles to penetrate skin (Di Cristo et al., in preparation).

Tab. 1: The predefined inhalation hypotheses

H-I-1	Respirable, biopersistent, rigid HARN: Following inhalation exposure, long-term pulmonary retention of HARNs can occur resulting in lung toxicity.	Murphy et al., 2021
H-I-2	Respirable, biopersistent, rigid HARN: Following inhalation exposure and translocation of HARNs to the pleura, mesothelioma development can occur.	Murphy et al., 2021
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.	Braakhuis et al., 2021
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.	Braakhuis et al., 2021
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.	Braakhuis et al., 2021
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.	Braakhuis et al., 2021

Tab. 2: The predefined oral-gastrointestinal hypotheses

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.	Di Cristo et al., 2021
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	Di Cristo et al., 2021
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.	Di Cristo et al., 2021
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.	Di Cristo et al., 2021
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.	Di Cristo et al., 2021
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.	Di Cristo et al., 2021
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.	Di Cristo et al., 2021
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 biopersistence in the body and systemic toxicity in secondary organs.	Di Cristo et al., 2021
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	Di Cristo et al., 2021

Tab. 3: The predefined dermal hypotheses

H-D-1	NFs with constituent substance(s) or degradation products classified for dermal irritation or sensitization: Dermal exposure to the NFs will result in comparable dermal irritation or sensitization depending on NF dissolution rate.	Di Cristo et al., in preparation
H-D-2	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 instantaneously releasing, dissolving and/or transforming into the same ionic or molecular forms.	Di Cristo et al., in preparation
H-D-3	NFs that are not biopersistent: Dermal exposure to NFs will not lead to accumulation of NFs or subsequent systemic toxicity.	Di Cristo et al., in preparation
H-D-4	NFs that are larger than 5nm and which are not flexible: Following dermal exposure NFs will result in limited or no dermal absorption and no dermal or systemic toxicity.	Di Cristo et al., in preparation

ALTEX, accepted manuscript
published June 22, 2022, version 3
doi:10.14573/altex.2203241

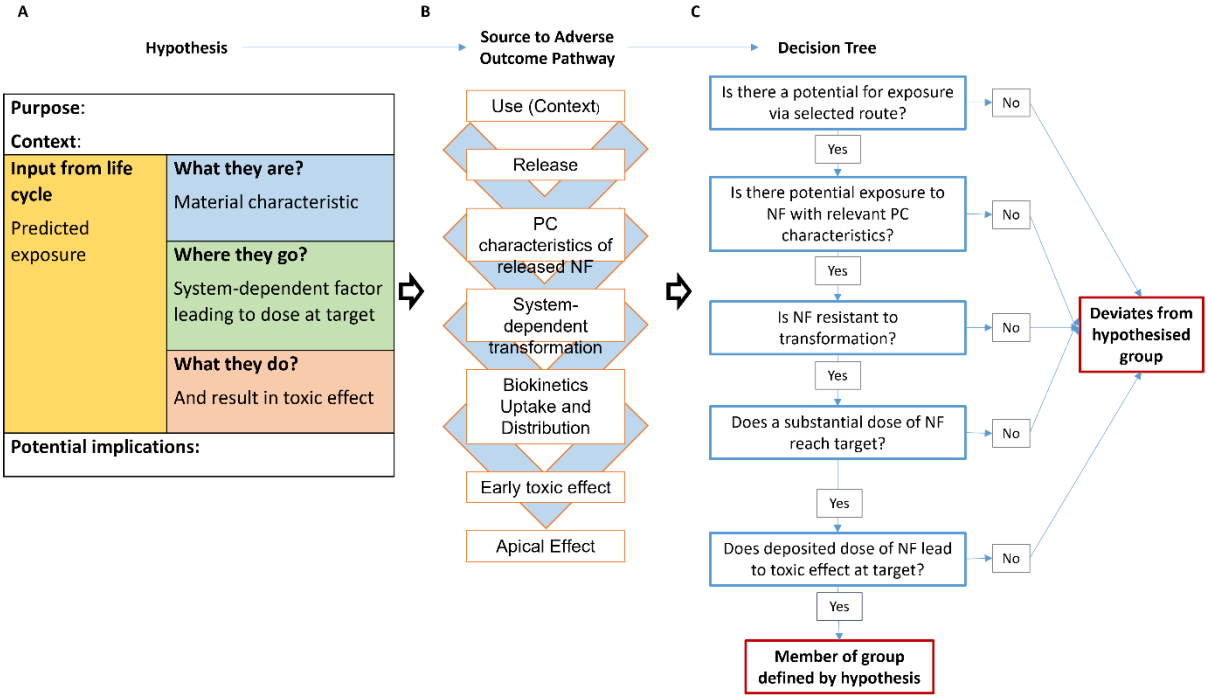

Fig. 2: Building a hypothesis-driven IATA
Information from hypothesis template (A) is converted into a Decision Tree (C) format following the concept of a source-to-adverse-outcome pathway (B).

The hazard endpoints identified in the hypotheses are varied, ranging from low specificity hazards such as 'lung toxicity' through to higher specificity endpoints such as mesothelioma development, reflecting the confidence and specificity of currently available structure/activity relationships for particle-driven hazards. Some of the hypotheses give an indication of the importance of the location where NFs might dissolve, especially for the NFs that do not instantly dissolve and can therefore translocate through cells and the body as particles before releasing constituent ions/molecules. This information also allows the hypotheses to consider both local effects relevant to the tissues exposed initially (including antimicrobial effects in the gut lumen) or systemically following translocation (e.g. liver) (Di Cristo et al., 2021). The dissolution and biopersistence information is also key to predict bioaccumulation and potential for longer-term hazards.

For the dermal hypotheses, links were also possible to the United Nation's Globally Harmonised System (GHS) (United Nations, 2021), which in the EU is implemented by the regulation on the classification, labelling and packaging of substances and mixtures (CLP Regulation) (EC, 2008), including effects such as skin sensitization (Di Cristo et al., in preparation).

5 Generating tailored IATA to support hypothesis testing

Structuring the hypothesis in the form of the template presented in Figure 1 can form the basis for generating a tailored IATA. The IATAs were designed to guide information gathering and the testing strategy required to make an evidence-based decision on whether the grouping hypothesis should be accepted or rejected for a specific group of NFs. The process of building an IATA from the generated hypothesis is summarized in Figure 2.

Once the hypothesis has been generated, a mechanistic pathway can be constructed detailing how the PC properties of the NF influence its movement along a proposed route from exposure to the target organ or environmental compartment for hazard assessment (Figure 2B). This is in line with the concept of a complete source-to-adverse-outcome pathway (SAOP) whereby all the different steps from first release to apical biological effect are considered relevant for NF grouping (Landsiedel, 2016; Oomen et al., 2015). This visual representation of the pathway from exposure to the target site allows the key PC characteristics directing fate and hazard to be identified, along with potential points of departure where the user may deviate from, and therefore reject, the grouping hypothesis (Figure 2C). The mechanistic links highlighted in the pathway are crucial to support the foundation of a grouping hypothesis and read-across argument, as the supporting data will serve to justify the premise of the group formation. If the hazard of the pristine material needs to be considered, e.g. for certain regulatory requirements, considerations on the exposure route are of relevance whereas the release and PC characteristics of released NFs are not necessarily taken into consideration.

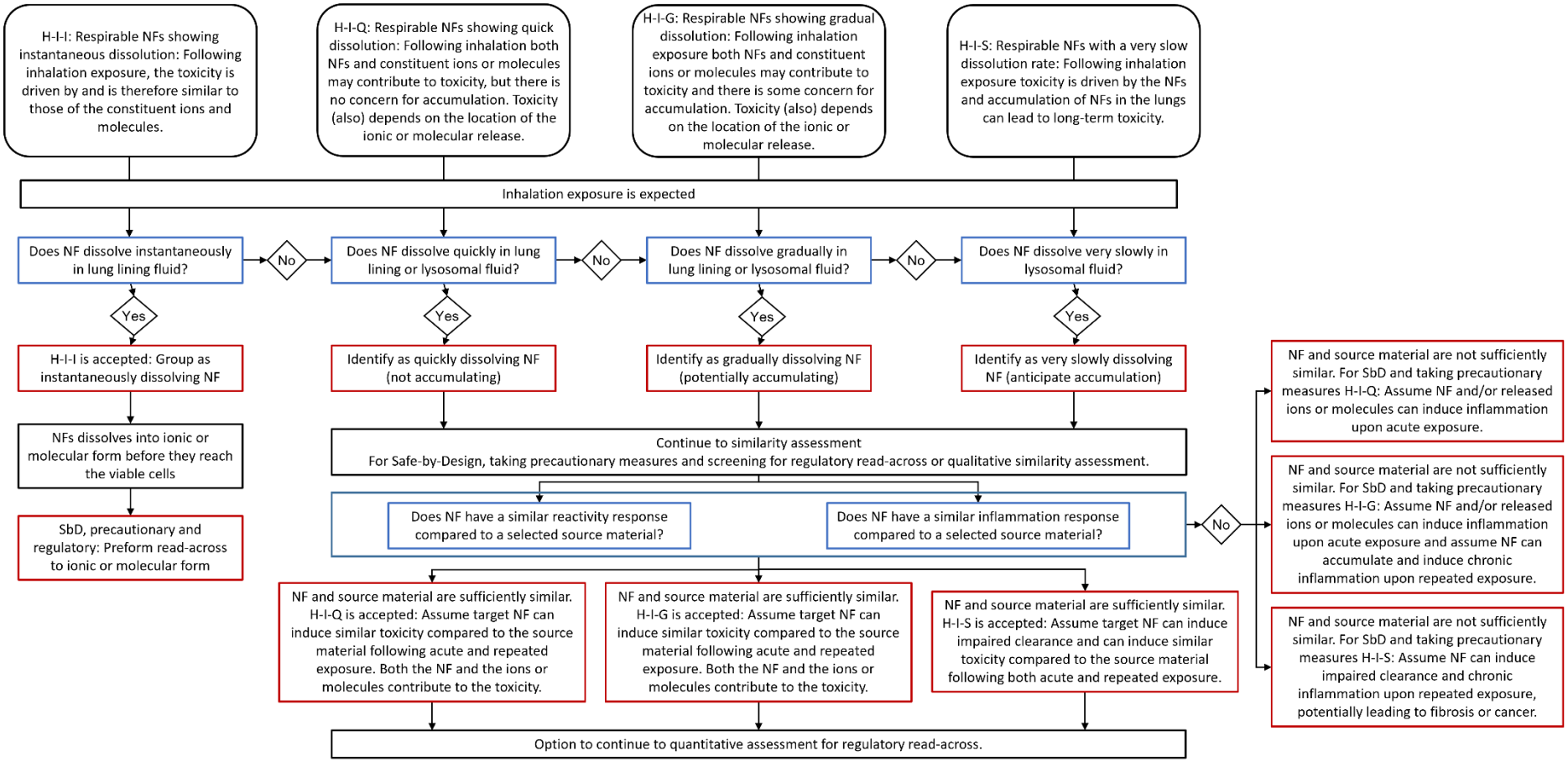

Fig. 3: The human inhalation IATA, adapted from Braakhuis et al., 2021
The remaining human health IATAs are provided in the supplementary information[6].

Identification of the critical intrinsic and system-dependent descriptors allows the pathway to take the form of a decision tree. More specifically, critical descriptors are incorporated as decision nodes controlling the progression through this pathway; this allows questions to be answered in order to determine whether the grouping hypothesis is valid for the NF under investigation (Figure 2). The evidence regarding the critical descriptors (and associated thresholds) needs to be scientifically underpinned and transparently described.

To decide whether the grouping hypothesis can be accepted or rejected, the IATA allows the integration of existing information (e.g., from peer-reviewed literature, reports and databases), modelling (i.e. *in silico* approaches) and experimental testing (fate, behavior, exposure, toxicokinetics and/or hazard) (OECD, 2017b). If generation of new data is required, the IATA provides recommendations for a practical and targeted testing strategy for collecting the required evidence through the identification of the most appropriate test methods currently available for each endpoint. Where possible, the critical descriptors forming decision nodes of the decision tree should be measurable, with defined thresholds to enable judgement on the similarity of the target NF to a source material selected from the group. The current state-of-the-art should be reviewed to identify relevant assays which provide the required data to answer the decision node questions, and utilize validated standard operating procedures, demonstrated as appropriate for testing of NMs.

6 The GRACIOUS IATAs

For every grouping hypothesis, the GRACIOUS project developed a tailored IATA, designed to gather only the information needed to test the grouping hypothesis. Such an approach should help to streamline data gathering and generation. Initially these IATAs were generated individually, one per hypothesis. As the IATAs evolved, common features and decision nodes were identified. Once the entire content of each IATA was established, the IATAs were aligned to identify common decision nodes and the logical flow of information gathering established. This process allowed the generation of combined IATAs for individual routes of exposure. The combined inhalation IATA which addresses non-HARN is presented in Figure 3.

Figure 3 is adapted from Braakhuis et al. (2021) and illustrates how the four relevant hypotheses require an initial consideration of the dissolution rate of the NFs. This includes dissolution in lung lining fluid (LLF) initially as the first compartment receiving the deposited particles. A tiered testing strategy is provided for this decision node which indicates the most relevant methods available to assess such dissolution (ISO, 2017). Assessment of dissolution rate in LLF allows grouping as either instantaneously, quickly, gradually or slowly dissolving. If dissolution is instantaneous, then the NF exposure can be considered equivalent to the inhalation exposure of the constituent ions or molecules. No further testing of the NF is required, and either read-across from existing data can be used, or new data for the ions or molecules can be generated. If dissolution is not instantaneous, then the user needs to consider assessing dissolution also in phagolysosomal fluid simulant at pH 4.5 to take into account potential for cellular uptake. Again, the protocols are provided via the tiered testing strategy. Both quickly and gradually dissolving NFs require a subsequent consideration of the hazards of both the solid NF and the released components, while the slowly dissolving group considers the NF particle alone.

A comparable structure to the IATA is provided for the oral route of exposure (Di Cristo et al., 2021), with LLF substituted for oral-gastrointestinal fluid simulants. Here, a protocol to simulate NF digestion is proposed. It allows the consecutive addition of simulant fluids (sequential addition of saliva, stomach and intestine fluids) varying in composition and pH to mimic the transit of the NF through the oro-gastrointestinal tract (Bove et al., 2017; Carnovale et al., 2021).

Due to the limited data sets available in the current literature, the dermal grouping IATAs are the least well developed. Nevertheless, it is clear that dissolution rate should be included as an initial decision node in most contexts to clarify whether the particulate NF and/or constituent ions and molecules are the likely drivers of potential hazard and therefore the most relevant for grouping. In order to address some of the more pertinent data gaps in the assessment of NF dissolution in the context of skin exposure, new data has been generated within the GRACIOUS project. The new data assesses the predictivity of simulated sweat fluids for NF dissolution (Di Cristo et al., in preparation), based on the preferred methodologies recommended within the inhalation and oral IATA tiered testing strategies, when appropriate. PSF is thus used to assess potential biopersistence of NFs that may cross the skin barrier and be taken up by Langerhans and macrophage cells in the underlying dermal layers.

Some of the hazard decision nodes for the inhalation, oral and dermal IATAs are also similar in content and structure, including cytotoxicity, surface reactivity, inflammation and genotoxicity. This commonality in decision node renders the IATA modular and demonstrates how decision nodes may be readily applicable to the construction of new IATAs. When the initial inhalation hypothesis was developed, insufficient information to support hypothesis generation for genotoxicity of NFs was identified. However, in order to future-proof the Framework, Verdon et al. (2022) have generated a genotoxicity tiered testing strategy that can be added into the existing IATA to address this gap, should future hypotheses require consideration of genotoxicity (Figure S1[6]).

In addition to supporting the gathering and generation of data, the IATAs also indicate where similarity assessments are required. In fact, a data matrix (Table S1[6]) can be generated that incorporates data for all tiers of every decision node for an individual IATA. In this way, a systematic analysis of the similarity of the data for each decision node can be generated. It is not possible to conduct a similarity assessment across data generated by different methods and hence across different tiers of the testing strategy. Instead, a similarity assessment can only be conducted for data generated by the same protocol. For this reason, use of standardized protocols is preferred. A description of the methods used to assess similarity are beyond the scope of this paper; they are however provided in a special issue of NanoImpact, summarized in the White paper by (Jeliazkova et al., 2022).

[6] doi:10.14573/altex.2203241s

Each of the IATAs presented have been tested using case studies. Both the inhalation IATA and the oral IATA have been tested with a panel of silica NFs (Braakhuis et al. in preparation, Di Cristo et al. in preparation), while the HARN IATA has been tested with diverse NFs of multi-walled carbon nanotubes (MWCNTs) (Murphy et al. in preparation). For the oral IATA, the case study concluded that the silica NFs were best described by a hypothesis for gradually dissolving NFs (Di Cristo et al., 2021). The original hypothesis indicates local toxicity and inflammation as the relevant hazard endpoint. However, the silica NFs did not generate any detectable local toxicity *in vitro* or *in vivo*, and so the case study-derived information allowed the hypothesis to be refined and made more specific to this group. The similarity assessment was successfully based upon Tier 1 *in vitro* methods, and further strengthened by the application of Tier 2 *in vitro* methods. Therefore read-across could be successfully used to fill Tier 3 data gaps and to conclude that all NFs within the group were gradually dissolving and did not induce any local toxicity in the gastrointestinal tract tissues. The conclusions for the hypothesis and IATA related to systemic effects were less consistent, and so similarity of the NFs in relation to distal effects was not sufficient to support grouping. The IATA therefore led to the rejection of the original grouping hypothesis, and the authors concluded that a case-by-case assessment of risk for these NFs with respect to systemic effects would be more appropriate at this time. The data matrix generated by this process is of course not wasted, as it can be recycled to identify data gaps and to structure the required individual risk assessment.

For the inhalation IATA case study, groups were formed that represented four groups of silica: quartz silicas (crystalline), precipitated amorphous silicas, pyrogenic amorphous silicas and colloidal silica NFs (Braakhuis et al., in preparation). This grouping could be achieved using Tier 1 assays for all decision nodes, and certainty was further enhanced by the inclusion of Tier 2 assays. All of these results were in line with available Tier 3 *in vivo* data, again supporting the suitability of the IATAs to support grouping and risk decision making.

The HARN inhalation case study used 15 different MWCNTs, with data obtained from the peer-reviewed literature (Murphy et al., in preparation). Application of the HARN IATA was followed by a pairwise similarity assessment of the grouped HARNs coupled with expert judgement. The IATA outcome allowed two of the MWCNT panel to be grouped according to the hypothesis and supported the application of read-across for pulmonary hazard endpoints between these two group members. The hypothesis indicated that inhalation of these respirable, biopersistent, rigid HARNs would lead to long-term pulmonary retention, resulting in lung toxicity characterized by inflammation, fibrosis and carcinogenicity. The second hypothesis that inhalation of respirable, biopersistent, rigid HARNs would result in translocation to the pleura and the subsequent development of mesothelioma was also supported. The case study concludes that the remaining 13 MWCNTs did not exhibit the appropriate physicochemical characteristics to be part of the group described by either of these hypotheses, and that instead the inhalation hypotheses for non-HARN particles proposed by (Braakhuis et al., 2021) should be considered for them.

Identification of dermal case studies have been more challenging due to the lack of available data.

7 Generation of new grouping hypotheses for future implementation in the GRACIOUS Framework

If the pre-defined hypotheses are not considered appropriate for a user's needs, the template can be used to develop a user-defined hypothesis. The user-defined hypothesis can be further developed into a pre-defined hypothesis for future inclusion in the GRACIOUS hypothesis tables.

Figure 4 outlines a roadmap for the development of a defined hypothesis for grouping. The figure illustrates that the full development process of a hypothesis proceeds in a stepwise fashion from the initial generation of the hypothesis using the template to the construction of IATA, followed by the acceptance, rejection or refinement of the hypothesis. The process of

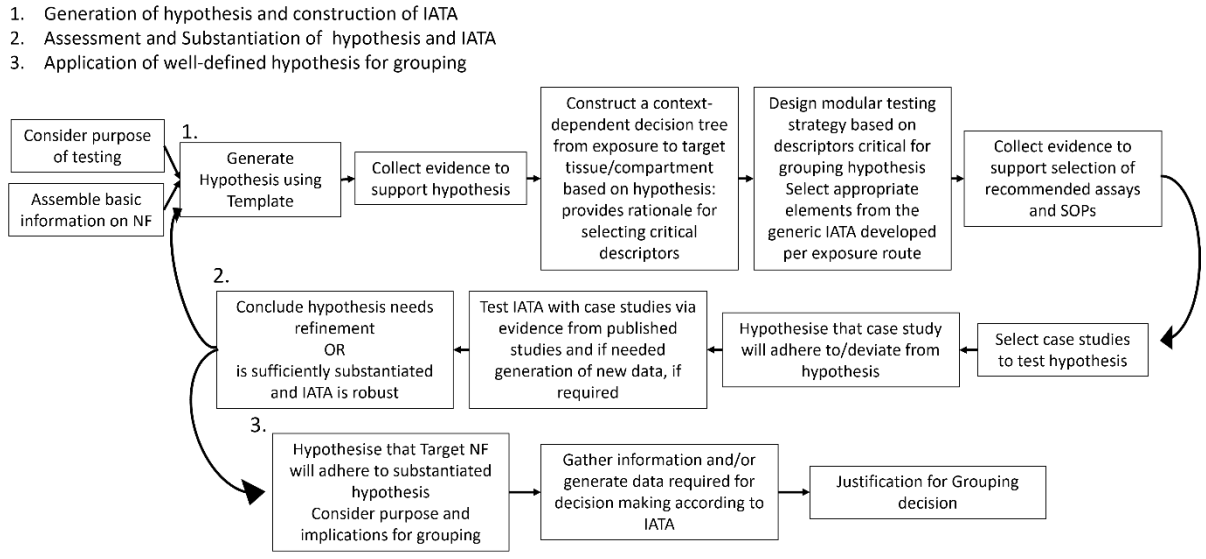

Fig. 4: Roadmap to the development of a substantiated hypothesis for grouping and read-across within the GRACIOUS framework

combining individual, hypothesis-specific IATAs facilitated by the alignment of common decision nodes (outlined in Section 6, Figure 3) nicely demonstrates the process by which a user may incorporate existing decision nodes into a user-specific IATA to test and support a newly-generated grouping hypothesis. Once a hypothesis has been developed, available data can be used to determine if the hypothesis is valid and applicable to the NF under investigation. It is recommended that the user follows defined criteria (Fernández-Cruz et al., 2018) for assessing the quality of the characterization of NFs in the studies included in the hypothesis assessment. Case studies should then be used to test the robustness of the hypothesis and the appropriateness of the associated IATA. At this point, a conclusion may be drawn, based on expert judgement, as to whether sufficient evidence has been collected to substantiate the hypothesis. If the evidence is considered insufficient or contradictory, refinement of the hypothesis is recommended. If the hypothesis requires refinement, the critical descriptors, boundaries, thresholds and data gaps described via the hypothesis template can be readily identified, enabling a well-structured refinement strategy to be adopted. Alternatively, if the hypothesis is substantiated and the IATA is considered appropriate, the hypothesis may be employed within the GRACIOUS framework for grouping or read-across for the target NF, with a sufficient level of confidence in the grouping decision considering the stated purpose of grouping.

New projects (e.g., SUNSHINE and HARMLESS) are currently employing the GRACIOUS Framework to develop grouping and read-across for multi-component nanomaterials. The approach taken anticipates the use of the existing hypotheses and IATAs where possible, potentially combining several to address the different components of the materials. It is also anticipated that additional controls will be required, as well as considerations of mixture effects and interactions (e.g., synergism). Such projects will be useful to expand the array of grouping hypotheses and IATAs available.

8 Progressing from grouping to read-across

Above, we have described how the IATA determine whether NFs (and non-NFs) are sufficiently similar with respect to a specific endpoint (e.g., dissolution) to be grouped. Once the group members are considered sufficiently similar and the grouping hypothesis is accepted, potential source material(s) with Tier 3 data need to be identified from within the group. These source materials will provide the data to be read across to the target NFs which lack the tier 3 data required for risk assessment.

(Stone et al., 2020) describes the specific example of instantaneously dissolving NFs, for which read-across from the solute for multiple endpoints is relevant. In the case studies described in the previous section, source materials were chosen based upon the availability of *in vivo* data, which for silica NFs was obtained for materials from the JRC Nanomaterials Repository[7] (JRCNM02000a and JRCNM02003a, also known as NM-200 and NM-203) and for HARNs included the Mitsui-7 MWCNT.

For read-across, the wording of the grouping hypothesis requires reformulation to focus on filling specific data gaps. This can be achieved by assessing whether the target NFs exhibit a similar hazard to the source NF(s) or non-NF(s). A suggested template for the generation of such a hypothesis is provided in the GRACIOUS Framework Guidance Document (Hunt, 2021) and is reproduced below.

***"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."*

9 Discussion

In this paper, we present a template to aid the construction of structured and evidence-based grouping hypotheses for NFs, and the resulting range of grouping hypotheses relevant to inhalation, oral and dermal exposure of humans. We also present the IATAs developed to support the targeted testing of the hypotheses.

The hypothesis template has been demonstrated to guide the user in the process of hypothesis generation for grouping of NFs. The initial step of using the hypothesis template defines the purpose for grouping, and this allows the user to determine the level of uncertainty seen as acceptable for their specific purpose. The adoption of the grouping hypothesis template by the wider nanosafety community will promote consistency in the design of grouping hypotheses and enable more rapid assessment and critique of the proposed grouping hypotheses, thereby helping to streamline the regulatory process in the future.

An additional advantage of the template is that it structures the hypothesis in a format that easily informs the design of a tailored IATA. The IATA can be used to generate the information needed to support a grouping decision relevant to the user's needs, thereby minimizing testing. A total of 19 hypotheses and complimentary IATAs relevant to human health, covering inhalation, oral and dermal routes of exposure, were generated using this template. The practical application and robustness of these hypotheses and IATAs have been tested by a number of internal case studies (completed as referenced in this paper or currently in progress) which demonstrated successful grouping of a number of panels of the NFs under investigation. Six additional case studies were also generated by stakeholders who were not part of the GRACIOUS consortium. One case study was led by the German Federal institute for Risk Assessment (BfR) and involved the RIVM, the German Federal Institute for Occupational Safety and Health (BAuA), the European Chemicals Agency (ECHA) and the European Commission represented by the Joint Research Centre (JRC). A second case study was conducted by the US National Institute of Occupational Safety and Health (NIOSH). Both of these case studies focused on MWCNT and independently verified some of the results of the GRACIUOS internal case study by Murphy et al. (under review). Both the NIOSH work

[7] https://ec.europa.eu/jrc/en/scientific-tool/jrc-nanomaterials-repository

and the Murphy et al. study identified the need to include length distribution data (rather than a mean value) and to better characterize the secondary structure of the MWCNTs, distinguishing between individual fibers, aligned fibrous bundles, granular agglomerates and tangled fibrous bundles. In a separate case study, Arche Consulting worked with the International Zinc Association to test an environmental hypothesis, focusing specifically on the dissolution decision node for selected soluble metallic NFs to understand if they can be grouped according to a high, intermediate or low toxicity ratio for solutes to particles. They applied time-weighted average concentrations to accurately express the exposure concentrations of the particle/ion mixture. They assumed toxicity between these components was additive, allowing quantification of the contributions of particles and ions to the overall suspension toxicity. They established a size dependent toxicity which was applicable to NFs of different composition, allowing read-across between NFs with different core chemical compositions. While read-across between NFs of different substances is not usually done in regulatory dossiers, such information can be useful for comparison to benchmark materials and also to support decision making in innovation. A number of confidential case studies were also conducted for industry via the consultant Blue Frog Scientific, resulting in advice on the clarity and level of detail required for testing.

Both internal and external case studies identified a number of issues which need to be considered when applying the template for real-life grouping, such as:

- Data availability - Data obtained from published literature was not always generated by the same method, and data gaps often occurred, limiting the ability to draw grouping conclusions.
- Aligning data sources - It was important to compare data generated via the same assay (and in some cases even the same assay conditions of a specific assay matter) when conducting grouping for regulatory purposes. If comparisons are made using different methods, only qualitative conclusions on grouping can be generated. Such qualitative assessments can still be useful, in particular when the purpose of the grouping is safe(r)-by-design or targeted testing (Braakhuis et al., 2021).
- Choice of assays to answer the questions in the decision nodes – for many in vitro assays, it is not demonstrated (yet) that they are sufficiently predictive for the in vivo situation. Preferably, simple Tier 1 assays are used to substantiate grouping, especially for the purpose of safe(r)-by-design. However, for some decision nodes, more complex Tier 2 or even Tier 3 assays might be needed to enhance predictivity. Further research is needed to evaluate which assays are most optimal to include for each decision node in an IATA.
- Dose metrics - Dose metrics affect the conclusions drawn. For example, for fibers, mean length metrics are typically reported in publications, while for substantiating a grouping decision, length size distributions are needed (Fraser et al., 2020). This was an important outcome independently reported by the internal MWCNT study and both external (i.e., BfR and US NIOSH) case studies. Another example is that NFs differ in reactivity and/or inflammatory potential when compared on a mass dose basis, but when doses are compared on a surface area basis, these differences are minimized. This finding is not novel (Duffin et al., 2007), but it was substantiated within GRACIOUS through the application of different similarity methods.
- Biological relevance of the outcome - Different assays may have different maximum responses, and the biological relevant range can vary. Again, this finding as such is not novel, but GRACIOUS substantiated it by applying various biostatistics/ bioinformatics methods in order to specifically define the biological useful dynamic range of specific methods.

We were unable to conduct dermal case studies due to the lack of available information for this route of exposure. This is a gap which will need to be addressed in future to ensure the suitability of the hypotheses and IATAs for grouping and read-across.

An evidence-based, hypothesis-driven approach is fundamental for a grouping or read-across argument to be considered acceptable within a regulatory context by agencies such as ECHA (ECHA, 2017a). ECHA provides guidance on building a read-across justification and evaluating a read-across case (ECHA, 2017a). However, the acceptance of read-across arguments is often hampered by industry partners submitting insufficiently justified arguments. This is for a large part due to lack of clarity regarding what regulators consider to be an acceptable level of uncertainty within the read-across argument (Patlewicz et al., 2015). (Ball et al., 2016) conducted a comprehensive review of the use of read-across in the REACH registration of non-nano substances and compiled the feedback received from ECHA on the acceptability of the read-across justifications. Based on their analysis of the ECHA compliance checks and review of testing proposals, the authors reported that a large proportion of the submitted read-across arguments have been rejected (Ball et al., 2016). The authors identified 107 compliance check decisions that involved the use of read-across, but only two were accepted, whereas 50 out of 81 testing proposals incorporating read-across were approved at least in some part. Many of the decisions highlighted that the approval of the testing plan should not be interpreted as approval of the read-across argument, and justification of the read-across was still required (Ball et al., 2016). Lack of supporting information and the use of unsubstantiated statements to form the basis of the hypothesis were identified as key reasons for rejection. Furthermore, questions on the scientific plausibility resulting from a combination of insufficient information and inadequate or conflicting information were also cited as grounds for rejection. Conversely, the successful applications of read-across were associated with proposals supported by robust justification that addressed all identified areas of uncertainty (Ball et al., 2016).

The presented hypothesis template, hypotheses and IATAs, when combined into the GRACIOUS Framework, provide the structure required for relevant grouping and read-across hypothesis development and substantiation. This approach highlights where evidence may be lacking for key statements related to PC characteristics, fate, toxicokinetics, hazard and the mechanistic links between these factors, which form the basis of the read-across argument; this in turn allows the data gaps and areas of uncertainty to be addressed. The hypothesis template therefore provides additional clarity to read-across arguments in line with RAAF guidance (ECHA, 2017a), which will facilitate improving the quality of read-across justifications submitted within registration dossiers of substances including NFs. The template may, in future, also be useful for justifying the grouping

of other types of chemicals and substances. The extended use of the template will provide regulators with a consistent and structured format against which to assess the abundance of read-across arguments currently being generated for NFs and chemicals alike. Recommendation of the use of a hypothesis template to support grouping and read-across could be an impactful addition to the RAAF and aid in streamlining the successful adoption and use of read-across in risk assessment.

The hypothesis template is designed to facilitate endpoint-specific grouping (in terms of common fate and/or common hazard potential between members of the group) rather than simply grouping based on shared PC characteristics. By directly posing the questions 'where they go' and 'what they do', the template readily highlights where there is a lack of sufficient information required to form a grouping hypothesis with clear implications relevant to purpose. Use of the hypothesis template to determine if there is sufficient information available to define a toxicologically relevant group promotes targeted testing to fill knowledge gaps, guiding users to the most relevant testing. Formation of a toxicologically relevant group is based on evidence that a causative relationship exists between the descriptors defining a group and a specific fate/hazard endpoint and is currently largely informed by known toxic modes of action triggered by exposure to pathogenic substances. Machine learning methods, e.g. Bayesian networks may be used to identify novel patterns of responses to NFs or drivers of toxicity leading to the formulation of new grouping hypotheses which may be supported by the generation of empirical evidence or statistical techniques if considered sufficiently robust. Case studies utilizing such computational methods should be conducted to demonstrate the potential application and limitations of these approaches.

A common ambiguity present in many (eco)toxicological studies concerns the insufficient characterization of the NF in realistic exposure environments (Krug, 2014). The linking of a hazard to PC properties refers to characterization carried out on pristine materials rather than the potentially transformed NF which may be released during various points of the NF lifecycle, and this may be misleading (Arts et al., 2016). For NFs that may have toxic effects, the predicted (eco)toxicity of a NF based on the PC characteristics of the pristine NF may differ significantly from that of the NF released in realistic exposure scenarios, due to the impact of extrinsic or system-dependent factors (Arts et al., 2014). The structured hypothesis approach described here will support the development of a functionality-driven approach to (eco)toxicity testing, encompassing the NF's intrinsic properties and its biophysical interactions with its environment in the context of the NF's life cycle.

Furthermore, the use of the grouping hypothesis template and IATAs can aid in the selection of appropriate source materials for comparison regarding hazard endpoints and toxicokinetics and support the development of a read-across argument. This manuscript identifies several case studies in which source materials have been successfully identified and used for both grouping and read-across, allowing acceptance of the original hypothesis (e.g., HARN inhalation), modification of the original hypothesis (e.g., local effects following oral and inhalation exposure to silica NFs) or rejection of the hypothesis (e.g., oral exposure to silica NFs for distal effects).

Our approach therefore provides a clear foundation from which a grouping decision can be supported. Further development of this approach will continue via the EU funded projects NanoRIGO, DIAGONAL, HARMLESS and SUNSHINE, allowing application to a wider and more complex array of materials. The OECD is also considering a future update of section 6.9 (on nanomaterials) of their guidance on the grouping of chemicals (OECD, 2017a). The GRACIOUS Framework, hypotheses and IATAs are contributing to the ongoing updating of this document, promoting wider use of the hypotheses, template, IATAs and Framework.

10 Conclusion

The template, hypotheses and IATAs presented here can facilitate both the initial grouping of NFs and subsequent read-across by setting out the purpose of grouping, the context being considered and the boundaries that define the group linked to specific fate and/or hazard endpoints, and by tailoring the information gathering required. The IATAs utilize fundamental knowledge on key release pathways, nano/eco-bio interactions and mechanisms of action compiled from the literature and integrated with data newly generated to fill data gaps. These IATAs will significantly impact on the future scientific developments in the areas of NF hazard and exposure assessment, with further elucidation of the mechanisms linking intrinsic and extrinsic PC descriptors to the release and exposure potential, toxicokinetics, environmental behavior and fate, and (eco)toxicological effects of NFs. Together the template, the hypotheses and IATAs, when utilized within the GRACIOUS Framework, will result in well-substantiated grouping of NFs for a number of applications spanning the fields of regulation, innovation, industry and academia.

References

Arts, J. H., Irfan, M. A., Keene, A. M. et al. (2016). Case studies putting the decision-making framework for the grouping and testing of nanomaterials (df4nanogrouping) into practice. *Regul Toxicol Pharmacol 76*, 234-261. doi:10.1016/j.yrtph.2015.11.020

Arts, J. H. E., Hadi, M., Keene, A. M. et al. (2014). A critical appraisal of existing concepts for the grouping of nanomaterials. *Regulatory Toxicology and Pharmacology 70*, 492-506. doi:10.1016/j.yrtph.2014.07.025

Aschberger, K., Christensen, F. M., Rasmussen, K. et al. (2016). Feasibility and challenges of human health risk assessment for engineered nanomaterials. In (eds.), *Engineered nanoparticles and the environment: Biophysicochemical processes and toxicity*. doi:10.1002/9781119275855.ch21

Ball, N., Cronin, M. T., Shen, J. et al. (2016). Toward good read-across practice (grap) guidance. *Altex 33*, 149-166. doi:10.14573/altex.1601251

Bernstein, D. M., Mast, R., Anderson, R. et al. (1994). An experimental approach to the evaluation of the biopersistence of respirable synthetic fibers and minerals. *Environmental Health Perspectives 102*, 15-18. doi:10.1289/ehp.94102s515

Bernstein, D. M., Morscheidt, C., de Meringo, A. et al. (1997). The biopersistence of fibres following inhalation and

intratracheal instillation exposure. *The Annals of Occupational Hygiene 41*, 224-230.

Bottero, J.-Y., Auffan, M., Borschnek, D. et al. (2015). Nanotechnology, global development in the frame of environmental risk forecasting. A necessity of interdisciplinary researches. *Comptes Rendus Geoscience 347*, 35-42. doi:10.1016/j.crte.2014.10.004

Bove, P., Malvindi, M. A., Kote, S. S. et al. (2017). Dissolution test for risk assessment of nanoparticles: A pilot study. *Nanoscale 9*, 6315-6326. doi:10.1039/c6nr08131b

Braakhuis, H. M., Murphy, F., Ma-Hock, L. et al. (2021). An integrated approach to testing and assessment to support grouping and read-across of nanomaterials after inhalation exposure. *Applied In Vitro Toxicology 7*, 112-128. doi:10.1089/aivt.2021.0009

Broaddus, V. C., Everitt, J. I., Black, B. and Kane, A. B. (2011). Non-neoplastic and neoplastic pleural endpoints following fiber exposure. *J Toxicol Environ Health B 14*, 153-178. doi:10.1080/10937404.2011.556049

Burden, N., Aschberger, K., Chaudhry, Q. et al. (2017). The 3rs as a framework to support a 21st century approach for nanosafety assessment. *Nano Today 12*, 10-13. doi:10.1016/j.nantod.2016.06.007

Carnovale, C., Guarnieri, D., Di Cristo, L. et al. (2021). Biotransformation of silver nanoparticles into oro-gastrointestinal tract by integrated in vitro testing assay: Generation of exposure-dependent physical descriptors for nanomaterial grouping. *Nanomaterials 11*, 1587.

Cogliano, V. J., Baan, R. A., Straif, K. et al. (2008). Use of mechanistic data in iarc evaluations. *Environmental and Molecular Mutagenesis 49*, 100-109. doi:10.1002/em.20370

Craighead, J. E. and Mossman, B. T. (1982). The pathogenesis of asbestos-associated diseases. *N Engl J Med 306*, 1446-1455. doi:10.1056/nejm198206173062403

Davis, J. M., Addison, J., Bolton, R. E. et al. (1986). The pathogenicity of long versus short fibre samples of amosite asbestos administered to rats by inhalation and intraperitoneal injection. *Br J Exp Pathol 67*, 415-430.

Dekkers, S., Oomen, A. G., Bleeker, E. A. J. et al. (2016). Towards a nanospecific approach for risk assessment. *Regulatory Toxicology and Pharmacology 80*, 46-59. doi:10.1016/j.yrtph.2016.05.037

Di Cristo, L., Oomen, A. G., Dekkers, S. et al. (2021). Grouping hypotheses and an integrated approach to testing and assessment of nanomaterials following oral ingestion. *Nanomaterials 11*, 2623.

Donaldson, K., Murphy, F. A., Duffin, R. et al. (2010). Asbestos, carbon nanotubes and the pleural mesothelium: A review of the hypothesis regarding the role of long fibre retention in the parietal pleura, inflammation and mesothelioma. *Part Fibre Toxicol 7*, 5. doi:10.1186/1743-8977-7-5

Duffin, R., Tran, L., Brown, D. et al. (2007). Proinflammogenic effects of low-toxicity and metal nanoparticles in vivo and in vitro: Highlighting the role of particle surface area and surface reactivity. *Inhalation Toxicology 19*, 849-856. doi:10.1080/08958370701479323

EC (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), establishing a european chemicals agency, amending directive 1999/45/ec and repealing council regulation (eec) no 793/93 and commission regulation (ec) no 1488/94 as well as council directive 76/769/eec and commission directives 91/155/eec, 93/67/eec, 93/105/ec and 2000/21/ec, o j l396, 2006. https://eur-lex.europa.eu/eli/reg/2006/1907

EC (2008). Regulation (ec) no 1272/2008 of the european parliament and of the council of 16 december 2008 on classification, labelling and packaging of substances and mixtures, amending and repealing directives 67/548/eec and 1999/45/ec, and amending regulation (ec) no 1907/2006. Oj l 353, 31.12.2008, p. 1–1355. https://eur-lex.europa.eu/eli/reg/2008/1272

EC (2010). Ec (2010). Directive 2010/63/eu of the european parliament and of the council of 22 september 2010 on the protection of animals used for scientific purposes text with eea relevance. Oj l 276 of 20.10.2010, p. 33—79. https://eur-lex.europa.eu/eli/dir/2010/63

EC (2018). Commission regulation (eu) amending regulation (ec) no 1907/2006 of the european parliament and of the council on the registration, evaluation, authorisation and restriction of chemicals (reach) as regards annexes i, iii,vi, vii, viii, ix, x, xi, and xii to address nanoforms of substances. Oj l 308, 4.12.2018, p. 1–20. https://eur-lex.europa.eu/eli/reg/2018/1881

ECHA (2008). Guidance on information requirements and chemical safety assessment. Chapter r.6: Qsars and grouping of chemicals. Guidance for the implementation of reach. European chemicals agency (echa). https://echa.europa.eu/guidance-documents/guidance-on-information-requirements-and-chemical-safety-assessment

ECHA (2017a). Read-across assessment framework (raaf). European chemicals agency (echa). doi:10.2823/619212

ECHA (2017b). 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 (echa). doi:10.2823/273911

Fernandez-Cruz, M. L., Hernandez-Moreno, D., Catalan, J. et al. (2018). Quality evaluation of human and environmental toxicity studies performed with nanomaterials - the guidenano approach. *Environmental Science: Nano 5*, 381-397. doi:10.1039/C7EN00716G

Fraser, K., Kodali, V., Yanamala, N. et al. (2020). Physicochemical characterization and genotoxicity of the broad class of carbon nanotubes and nanofibers used or produced in u.S. Facilities. *Particle and Fibre Toxicology 17*, 62. doi:10.1186/s12989-020-00392-w

Grosse, Y., Loomis, D., Guyton, K. Z. et al. (2014). Carcinogenicity of fluoro-edenite, silicon carbide fibres and whiskers, and carbon nanotubes. *The Lancet Oncology 15*, 1427-1428.

Hunt, N. (2021). Guidance on the gracious framework for grouping and read-across of nanomaterials and nanoforms. doi:doi:10.5281/zenodo.5534465

ISO (2017). Iso/tr 19057:2017 nanotechnologies. Use and application of acellular in vitro tests and methodologies to assess nanomaterial biodurability. ISO/TR 19057:2017. https://www.iso.org/standard/63836.html

Jeliazkova, N., Bleeker, E., Cross, R. et al. (2022). How can we justify grouping of nanoforms for hazard assessment? Concepts

and tools to quantify similarity. *NanoImpact 25*, 100366. doi:doi:10.1016/j.impact.2021.100366
Krug, H. F. (2014). Nanosafety research--are we on the right track? *Angew Chem Int Ed Engl 53*, 12304-12319. doi:10.1002/anie.201403367
Kuempel, E. D., Geraci, C. L. and Schulte, P. A. (2012). Risk assessment and risk management of nanomaterials in the workplace: Translating research to practice. *The Annals of Occupational Hygiene 56*, 491-505. doi:10.1093/annhyg/mes040
Lamon, L., Aschberger, K., Asturiol, D. et al. (2019). Grouping of nanomaterials to read-across hazard endpoints: A review. *Nanotoxicology 13*, 100-118. doi:10.1080/17435390.2018.1506060
Landsiedel, R. (2016). Concern-driven integrated approaches for the grouping, testing and assessment of nanomaterials. *Environ Pollut 218*, 1376-1380. doi:10.1016/j.envpol.2015.10.060
Mech, A., Rasmussen, K., Jantunen, P. et al. (2019). Insights into possibilities for grouping and read-across for nanomaterials in eu chemicals legislation. *Nanotoxicology 13*, 119-141. doi:10.1080/17435390.2018.1513092
Murphy, F., Dekkers, S., Braakhuis, H. et al. (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:doi:10.1016/j.impact.2021.100314
Nagai, H., Okazaki, Y., Chew, S. H. et al. (2011). Diameter and rigidity of multiwalled carbon nanotubes are critical factors in mesothelial injury and carcinogenesis. *Proceedings of the National Academy of Sciences of the United States of America 108*, E1330-E1338. doi:10.1073/pnas.1110013108
OECD (2016). Grouping and read-across for the hazard assessment of manufactured nanomaterials. OECD Series on the Safety of Manufactured Nanomaterials. *OECD Publishing*, https://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2016)59&doclanguage=en
OECD (2017a). Guidance on grouping of chemicals, second edition. OECD Series on Testing and Assessment,. *P. OECD Publishing*, doi:10.1787/9789264274679-en
OECD (2017b). *Guidance document on the reporting of defined approaches to be used within integrated approaches to testing and assessment*. Vol. doi:10.1787/9789264274822-en
OECD (2018). *Test no. 451: Carcinogenicity studies*. Vol. doi:10.1787/9789264071186-en
Oomen, A. G., Bleeker, E. A., Bos, P. M. et al. (2015). Grouping and read-across approaches for risk assessment of nanomaterials. *Int J Environ Res Public Health 12*, 13415-13434. doi:10.3390/ijerph121013415
Patlewicz, G., Ball, N., Boogaard, P. J. et al. (2015). Building scientific confidence in the development and evaluation of read-across. *Regul Toxicol Pharmacol 72*, 117-133. doi:10.1016/j.yrtph.2015.03.015
Poland, C. A., Duffin, R., Kinloch, I. et al. (2008). Carbon nanotubes introduced into the abdominal cavity of mice show asbestos-like pathogenicity in a pilot study. *Nature Nanotechnology 3*, 423-428. doi:10.1038/nnano.2008.111
Rittinghausen, S., Hackbarth, A., Creutzenberg, O. et al. (2014). The carcinogenic effect of various multi-walled carbon nanotubes (mwcnts) after intraperitoneal injection in rats. *Part Fibre Toxicol 11*, 59. doi:10.1186/s12989-014-0059-z
Selikoff, I. J., Churg, J. and Hammond, E. C. (1964). Asbestos exposure and neoplasia. *JAMA 188*, 22-26. doi:10.1001/jama.1964.03060270028006
Sellers, K., Deleebeeck, N., Messiaen, M. et al. (2015). *Grouping nanomaterials - a strategy towards grouping and read-across*. Vol.
Stayner, L., Welch, L. S. and Lemen, R. (2013). The worldwide pandemic of asbestos-related diseases. *Annu Rev Public Health 34*, 205-216. doi:10.1146/annurev-publhealth-031811-124704
Stone, V., Gottardo, S., Bleeker, E. A. J. et al. (2020). A framework for grouping and read-across of nanomaterials- supporting innovation and risk assessment. *Nano Today 35*, 100941. doi:10.1016/j.nantod.2020.100941
United Nations (2021). *Globally Harmonized System of Classification and Labelling of Chemicals (GHS)*, 9th edition. https://unece.org/sites/default/files/2021-09/GHS_Rev9E_0.pdf
Verdon, R., Stone, V., Murphy, F. et al. (2022). The application of existing genotoxicity methodologies for grouping of nanomaterials: Towards an integrated approach to testing and assessment. *Particle and Fibre Toxicology 19*, 32. doi:10.1186/s12989-022-00476-9
Wagner, J. C., Sleggs, C. A. and Marchand, P. (1960). Diffuse pleural mesothelioma and asbestos exposure in the North Western Cape Province. *Br J Ind Med 17*, 260-271. doi:10.1136/oem.17.4.260

Conflict of interest

The authors have no conflicts of interest to declare.

Acknowledgements

The GRACIOUS project was supported by the European Commission Horizon 2020 programme, Grant Agreement No. 760840. The SUNSHINE project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 952924.

Data availability statement

No datasets were generated or analyzed during this study.