[1]Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

The nanoGRAVUR framework to group (nano)materials for their occupational, consumer, environmental risks based on a harmonized set of material properties, applied to 34 case studies†

Wendel Wohlleben,[b] Bryan Hellack,[g] Carmen Nickel,[g] Monika Herrchen,[e] Kerstin Hund-Rinke,[e] Katja Kettler,[c] Christian Riebeling,[c] Andrea Haase,[c] Birgit Funk,[i] Dana Kühnel[k], Daniel Göhler,[h] Michael Stintz,[h] Christian Schumacher,[d] Martin Wiemann,[f] Johannes Keller,[b] Robert Landsiedel,[b] Dirk Broßell,[a] Sabine Pitzko [a] and Thomas Kuhlbusch [a,j*]

The project nanoGRAVUR (BMBF, 2015-2018) developed a framework for grouping of nanomaterials. Different groups may result for each of the three distinct perspectives of occupational, consumer and environmental safety. The properties, methods and descriptors are harmonised between the three perspectives and are based on:

- Tier 1 intrinsic physico-chemical properties (what they are) or GHS classification of the non-nano-form (human tox, ecotox, physical hazards);
- Tier 2 extrinsic physico-chemical properties, release from nano-enabled products, *in vitro* assays with cells (where they go; what they do);
- Tier 3 case-specific tests, potentially *in vivo* studies to substantiate the similarity within groups or application-specific exposure testing

Amongst all properties, dissolution and transformation are least modulated by different nanoforms within one substance, whereas dustiness, dispersion stability, abiotic and especially *in vitro* surface reactivity vary more often between different nanoforms. The methods developed or selected by nanoGRAVUR fill several gaps highlighted in the ProSafe reviews, and are useful to implement i) the concept of nanoforms of the European Chemicals Agency (ECHA) and ii) the concept of discrete forms of the United States Environmental Protection Agency (EPA). One cannot assess the significance of a dissimilarity, if the dynamic range of that property is unknown. Benchmark materials span dynamic ranges that enable us to establish bands, often with order-of-magnitude ranges. In 34 case studies we observed high biological similarity within each substance when we compared different (nano)forms of SiO_2, $BaSO_4$, kaolin, CeO_2, ZnO, organic pigments, especially when we compared forms that are all untreated on the surface. In contrast, different Fe_2O_3 or TiO_2 (nano)forms differ more significantly. The same nanoforms were also integrated in nano-enabled products (NEPs) for automotive coatings, clinker-reduced cements, cosmetic sunscreen, and lightweight polymers.

[a.] *Federal Institute for Occupational Safety and Health (BAuA), 44149 Dortmund, Germany.*
[b.] *BASF SE, Dept. of Material Physics and Dept. of Experimental Toxicology and Ecology, 67056 Ludwigshafen, Germany.*
[c.] *Federal Institute for Risk Assessment - BfR. 10589, Berlin, Germany*
[d.] *German Social Accident Insurance e. V. -DGUV, 10117, Berlin, Germany*
[e.] *Fraunhofer Institute for Molecular Biology and Applied Ecology, 57392, Schmallenberg, Germany*
[f.] *IBE R&D Institute for Lung Health gGmbH, 48149 Münster, Germany*
[g.] *Institute of Energy and Environmental Technology e. V. – IUTA, 47229, Duisburg, Germany*
[h.] *Research Group Mechanical Process Engineering, Institute of Process Engineering and Environmental Technology, Technische Universität Dresden, D-01062 Dresden, Germany*
[i.] *Zoz GmbH, 57482, Wenden, Germany*
[j.] *CENIDE, University Duisburg-Essen, Germany*
[k.] *Helmholtz Centre for Environmental Research GmbH - UFZ, Leipzig, Germany*
* corresponding author, Kuhlbusch.Thomas@baua.bund.de
†Electronic Supplementary Information (ESI) available. See DOI: 10.1039/x0xx00000x

Introduction

Particles, i.e., minute pieces of matter with defined physical boundaries [2], are commercially available in a myriad of grades that are optimised in composition, size, shape and coating for specific applications.[3, 4] If the number metric median diameter of the constituent particles is below 100 nm, that grade is identified as nanomaterial for regulatory purposes by the definition that was recommended by the European Commission.[5, 6] The identification as nanomaterial is explicitly intended to be without regard to hazard or risk[5, 6], but triggers additional or more specific testing requirements from 2020 via the revised REACH Annexes.[7] Several frameworks have been proposed that structure the risk assessment of nanomaterials in tiered testing strategies,[8] often supported by elements of grouping and read-across strategies,[9, 10] often targeting a reduction of animal testing by alternative methods.[11-13] The tiered testing strategies serve two main purposes:

a. Industry has an interest to ensure the safe use of novel nanomaterials early during development, using a minimum amount (milligrams) of available material and without animal testing. If the safety testing feeds back to optimize the balance of performance, safety, costs and sustainability (or to stop the development), then the term "safer-by-design" is often used to describe this good industrial practice. The comparison of a novel nanomaterial to benchmark materials with a well-known (eco)toxicological profile can support the confidence in the assessment, but it is unknown which material properties should be used to make that comparison.
b. European regulators have established the concept of nanoforms (NF, nanomaterial form of a chemical which is characterised by ranges of morphology, particle size distribution, surface chemistry, specific surface area) to register forms of a substance that are identified as nanomaterial in the substance dossier.[14]. The data requirements in the revised REACH Annexes VI to X have been amended accordingly.[5] Concepts of similarity and grouping are relevant since „sets of similar nanoforms" can be registered, with a justification.[5, 15] Registrants can use concept of grouping for such justification and can use read-across to fill data gaps of one NF (or one set) by existing data of the non-nano-form or of another NF (or another set) of the same substance, if there is a specific hypothesis why "source" and "target" form should be similar.[16]. Although the guidance proposes specific properties to substantiate the similarity, it does not mark them as mandatory and lists no methods, nor descriptors, nor benchmark materials.[16].

There are numerous deficiencies of the existing regulatory guidances and of the published frameworks:

Most of the grouping frameworks, including the seminal NIOSH proposal [17] and its implementation [18], the DF4nanogrouping framework [19] and its implementations [20-22], limit their scope to human safety, more specifically to inhalation hazards. Regarding the prediction of the environmental hazard of nanomaterials two approaches are pursued. Besides modelling and the development of structure-activity-relationships [23], only a single framework assesses the environmental hazard from a grouping perspective with proposed trigger values for the identified physico-chemical properties to be relevant for ecotoxicity of metals and metal oxides [24, 25]. Several frameworks recognize that the integration of nanomaterials into nano-enabled products (NEPs) determines the material properties of the fragments that may be released throughout the life cycle [26] and screening e-tools such as the LICARA nanoscan [27], GuideNano,[28] SUN, [29, 30] select some relevant NEP properties, but no grouping framework integrates the safety of consumer use of NEPs or the safety of professional handling of NEPs.[8]

The DF4nanogrouping [19, 20] selects specific methods of analysis, quantitative cut-offs and benchmark materials.[8] In contrast, the ECHA nanomaterial grouping guidance refers to the generic physico-chemical guidance [31] that lists numerous optional methods, and refers also to the DF4nanogrouping method selection Table S2 and S4.[16] Often the frameworks used the same terminology and some select the same material properties to compare (nano)-forms,[8] but many properties may be determined by different analytical methods, and for some the reliability and standardisation is insufficient.[32] Additionally, there are numerous options of data reduction from the multidimensional raw data (images, spectra, distributions) to simple scalar descriptors (one numerical value) that can quantify the similarity between (nano)forms or the homogeneity of groups.

The project nanoGRAVUR (2015-2018) was funded by the German Federal Ministry of Research, and by industry and comprised partners from academia, regulatory agencies, insurance companies and industry. This paper presents the nanoGRAVUR grouping framework for nanomaterials and its implementation by selected methods of analysis and quantitative benchmark material values. Different groups may result for each of the three distinct perspectives of *Occupational, Consumer* and *Environmental safety (OCE)*, but rely on a harmonised set of material properties with specific methods of analysis, descriptors and ranges. The proof of concept is provided via quantitative data on 34 case studies.

Framework and selection of properties

The nanoGRAVUR grouping framework (Figure 1) consists of three tiers. Tier 1 determines intrinsic physicochemical properties ("what they are") and/or the GHS classification of the non-nano form (human tox, ecotox, physical hazards). Tier 1 allows the user to describe concerns and accordingly a grouping hypothesis. Depending on the purpose of grouping (Table 1), Tier 2 determines extrinsic physicochemical properties, the release from nano-enabled products (NEPs, if in scope) and/or *in vitro* assays ("what is the NEP"; "where they go"; "what they do"). If the assessment remains inconclusive, Tier 3 deals with case-specific testing, potentially *in vivo* studies to substantiate the similarity within groups or application-specific exposure testing.

The nanoGRAVUR grouping framework serves three purposes of testing and grouping, that each require a different perspective: occupational safety, environmental safety, consumer safety. Aspects of both hazard and exposure are considered, and the intended use in NEPs is systematically integrated in the testing strategy. Depending on the purpose of grouping, only a sub-selection of the material properties is required. The assessment is guided by

i. the purpose of grouping (Table 1)
ii. tiers with increasing specificity of testing (Figure 1)

Compared to the properties that are proposed (but not mandatory) by the ECHA guidance,[16], a core set of properties overlaps, some properties were added, and some were tested but not selected. Despite this difference, we view the nanoGRAVUR framework as an implementation of the ECHA guidance: If the purpose of grouping is to assess the risk of occupational handling of powders („O" in Table 1), our choice of properties is a *selection* of the properties proposed by ECHA, without additions. The reason for *selecting* is that we do not group by *intrinsic* properties that are only a proxy of biological interactions. This concerns especially properties that describe surface chemistry including hydrophobicity and charge. Instead, we group by *extrinsic* properties (dispersibility, biological reactivity, dissolution in biological and environmental media), that assess the interactions more directly, and correlate to human and ecological hazard and fate results.[33-35] The additional properties decrease the uncertainty of risk estimation at limited additional efforts, as they are descriptive (e.g. NEP categories / intended uses), rely on simple methods of analysis (see next section), and widen the scope to NEPs as relevant for the value chains related to professionals, consumers and the environment. While the NEP does not need to be included in the assessment, at least the intended use of the engineered nanomaterial (ENM) needs to be known to select a „relevant medium"in Tier 2. This reasoning was already implemented in the DF4nanogrouping,[19] but is expanded here to environmental media.

In the stepwise methodology of the ECHA grouping guidance, the nanoGRAVUR Tier 1 data requirements of „Primary particle shape", „Composition (incl. impurities)" with the corresponding „GHS/CLP human toxicity (bulk)", „GHS/CLP ecotoxicity (bulk)" and „GHS/CLP physical hazards (bulk)" constitutes ECHA Step 1. We derive our hypothesis of most important hazards via the data reduction to quantitative property ranges as a simple implementation of ECHA steps 2 to 4. The only grouping decision that can be taken after Tier 1 is the grouping of a highly soluble NF with their non-nano-form of the same substance, based on water solubility with a specific method of analysis and cut-off value (Table 2). All other cases will require testing by Tier 2 methods to substantiate the similarity of different NFs in a group. The testing in Tier 2 constitutes ECHA Steps 5 to 6, and the focus on extrinsic properties is supported explicitly by the ECHA guidance, reading „*It should be noted that differences in the physical parameters seen when characterising the nanoforms does not per se exclude the possibility to apply read-across. Indeed, similarities in the parameters related to the behaviour (e.g. solubility) or those relating to their reactivity may be more important to consider when building a read-across justification*".[16] The Tier 1_NEP data requirement of „NEP classes & intended use scenarios" adds a grouping hypothesis by relevant release and exposure pathways, and is essential to select relevant media in Tier 2. Of note, the same approach is taken by the US-Canada Regulatory Cooperation Council, where the prioritization of either of the human exposure pathways trigger dissolution testing in each different media, complemented by composition ionic toxicity, shape, and surface reactivity.[36]

The nanoGRAVUR scheme for ecological hazard grouping requires only Tier 2 results, and is consistent with the previous grouping developed on the basis of tests with algae, daphnia and fish embryo, that focused on ion release, reactivity, and shape [24] and its further development [25]. The additional property "surface affinity" was added, and further adaptations are not excluded. Grouping regarding environmental fate was based on a preliminary description of processes. Depending on the complexity of the studies, these are Tier 2 (e.g. dissolution in environmental media) or Tier 3 tests (e.g. mobility in soil and biological transformation). For the latter, no triggers were identified but these tests need to be performed to enable fate grouping.[25]. The proposed methods of analysis and some quantitative cut-off values are listed in Table 2.

Analogously, if the validity of the hypothesis for grouping by occupational safety remains inconclusive after Tier 2, specific testing in Tier 3 is possible. This may involve exposure testing at specific workplaces and may involve *in vivo* animal studies.

In Tier 2, the assessment of exposure is approximated by dustiness and NEP properties that are most relevant for release or emission. This enables risk-based groupings, but still supports hazard-based groupings as prioritised in REACH. Thus, a grouping hypothesis might be that different NEPs that have similar matrices with different embedded ENMs are similar in the rate and form of released fragments, and thus also similar in the hazard by such fragments, as motivated by results on human hazard [37-42] and ecological hazard [26, 43] of such fragments. Specific methods support the assessment of emission, fate and transport behaviour, as detailed in the following chapter, and thus expand from the groupings perspective of environmental hazard [24] and human hazard [19].

Framework concept in comparison to other approaches

There are fundamentally different grouping approaches. E.g., one may group not by measurement of material properties but by hazard testing. NIOSH grouped by *in vivo* potency.[18] This is equivalent to skipping Tiers 1 and 2, and performing only Tier 3 *in vivo*. However, the resulting groups have no logical relationship, as is provided by the intrinsic and extrinsic properties, and thus also deviates from the seminal NIOSH proposal of four groups delimited

by shape, solubility, and bulk toxicity [44] that was the basis of the AGS BekGS527 [45] and of the DF4nanogrouping[19], and is still recognisable in the nanoGRAVUR scheme. Another grouping by toxicity was performed via cytokine profiling,[46] but the experimental effort to generate such data may prevent a robust and pragmatic regulatory use, and may be more appropriate for mechanistic studies. However, it is interesting to note that both the grouping by *in vivo* potency and the grouping by cytokine profiling result in groups given primarily by the substance composition, e.g. grouping TiO_2 NFs, ZnO NFs and non-nano-forms, SiO_2 NFs.[18, 46] Yet another complementary categorisation was proposed by the FutureNanoNeeds project, combining by „information multiplexing" the physico-chemical properties, descriptors of the adsorbed corona (in an unspecified medium), and a mapping of the biologically accessible epitopes,[47] and the chemical basis of nano-bio interactions was thought to enable structure-activity-relationships.[48] With the sparse data on corona and epitopes, these concepts may be true but remain speculative. If validated, the FutureNanoNeeds concept might be an approach to group especially pharmaceutical nano-fomulations by their systemic transport, but the specificity would fall short of the Tier 3 requirements of specific uptake and clearance tests, and the experimental workload might not be justified against an *in vivo* study as valid option. We do not consider pharmaceutical purposes here.

Table 1 nanoGRAVUR selection of material properties. The set is harmonised across the three purposes of grouping for (O) Occupational, (C) Consumer and (E) Environmental risks, includes the risks in the value chain of nano-

	Properties (continued on next page)	nanoGRAVUR data in case studies	nanoGRAVUR (O) Occupation (C) Consumer (E) Environment	nanoGRAVUR Method TRL	EPA discrete forms (2017)	ECHA nanoform registration, best practice (2017)	ECHA grouping guidance (2017)	DF4nanogrouping framework (2015)	Environmental hazard grouping framework (2018)
What they are	Primary particle shape	✓	Criterion O,C,E	+	Criterion	Criterion	Proposed	Criterion O	Criterion E
	Primary particle dimension	✓	Criterion O,C,E	+	Criterion	Criterion	Proposed		Criterion E
	Composition + GHS (incl. Impurities)	✓	Criterion O,C,E	+	Criterion	Criterion	Proposed	Criterion O	Criterion E
	Specific surface area (BET/VSSA)	✓		+	Criterion		Proposed	Suppl. O	
	Surface Chemistry (descriptive)	✓		0	Criterion	Criterion	Proposed	Suppl. O	
	Surface Charge (zeta-potential)	✓		+	Criterion		Proposed	Suppl. O	
	Hydrophobicity	✓		-			Proposed	Suppl. O	
	Rigidity (for fibers)		Proposed O	-					
	NEP classes & intended use scenarios	(✓)	Criterion (NEP) O,C						
	Specific NEP: g/g content of ENM	✓	Criterion (NEP) O	+					
	Specific NEP: dispersion state of ENM	(✓)	Criterion (NEP) O,C	O					
	Properties (continued from previous page)	nanoGRAVUR data in case studies	nanoGRAVUR (O) occupation (C) Consumer (E) Environment	nanoGRAVUR Method TRL	EPA discrete forms (2017)	ECHA nanoform registration, best practice (2017)	ECHA grouping guidance (2017)	DF4nanogrouping framework (2015)	Environmental hazard grouping framework (2018)
Where they go: Release / Exposure	Dustiness	✓	Criterion O	+			Proposed	Qualifier O	
	Critical shapes upon exposure	(✓)	Criterion (NEP) O,C	0					
	Agglo. of ENM upon NEP application	(✓)	Criterion (NEP) O,C	0					
	Resilience of NEP Matrix	✓	Criterion (NEP) O,C	+					
Where they go: In relevant media	Dispersability (dispersion stability)	✓	Criterion O,C,E	+0			Proposed	Criterion O	
	Solubility in water (screening test)	✓	Criterion O,C,E	+			Proposed	Criterion O	
	Dissolution rate in relevant media	✓	Criterion O,C,E	0			Proposed	Criterion O	
	Ion releasing	(✓)	Criterion E	+0					Criterion E
	Transformation „change of what they are"	✓	Criterion E (Tier 3)	0-					
	Mobility (in soils)	✓	Criterion E (Tier 3)	+					
	Mobility (systemic) by altern. Method		Criterion C	-					
	Affinity (heteroagglomeration)		Proposed E	0					
What they do	Reactivity (abiotic)	✓	Criterion O,C,E	+0			Proposed	Criterion O	Criterion E
	Reactivity (in-vitro)	✓	Criterion O,C	+0				Criterion O	
	Reactivity (photo-)			-			Proposed		

enabled products (NEP), and can be compared to regulatory guidance and previous frameworks that addressed selected purposes.[1, 14, 16, 19, 24] Proposed: Property is not mandatory, but proposed for decision-making.; Criterion: Property with quantitative cut-off for decision-making; Supplementary: Property without use in decision-making; Qualifier: Required to select appropriate conditions in further testing. The symbols for maturity of the methods are indicative of Technology Readiness Level (TRL) 4 or lower (-), 5 to 7 (0), 8 or higher (+).

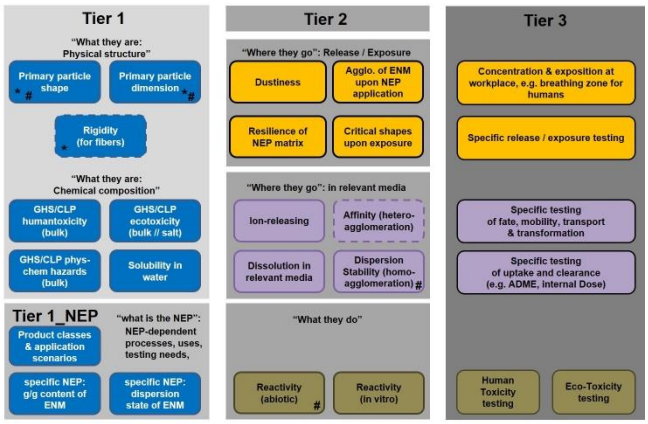

Figure 1 nanoGRAVUR grouping framework. ENM: engineered nanomaterial; NEP: nano-enabled product; GHS: globally harmonised system; CLP: Classification and Labelling of Products; * denotes properties used by ECHA to differentiate nanoforms;[14] # denotes properties used by EPA to differentiate discrete forms.[1] Dashed boxes denote properties with no sufficiently developed method.

Recommended methods to determine material properties

The selection of specific methods of analysis or combinations of methods is an integral element of the nanoGRAVUR grouping-framework. Table 2 specifies the selected methods of analysis, the data reduction to descriptors as well as recommended metrics, and representative benchmark materials (OECD NMs). Additional details on the SOPs are provided in the SI.

We harmonised the selection of methods across all different purposes of grouping, aiming to simplify the practical implementation, to reduce costs, and to maximise the multiple use of data. The ProSafe review on the reliability of methods for the regulatory assessment of nanomaterials indicated for each property several „preferred methods“, but also discussed their limitations and knowledge gaps.[32] The overlap between that ProSafe preference and the nanoGRAVUR selection is substantial, and supports the robustness of the nanoGRAVUR grouping framework. The nanoGRAVUR method selection deviates rarely from the ProSafe preference, and fills some essential knowledge gaps, as discussed in the following.

Tier 1 Methods for “what they are” properties

The methods and descriptors for „what they are“ properties concerning particle size distribution, shape, and specific surface area are implemented according to the recommendation of the project NanoDefine for methods (TEM and BET for powders, analytical centrifuges for suspensions), and descriptors (number metrics median).[50, 51] The sample preparation is an essential element and was included in the NanoDefine validation.[52] For the chemical composition we agreed with ProSafe that ICPMS is suitable, but recommend XRF as digestion-free proxy for composition and all inorganic impurities. Properties that were not selected for the nanoGRAVUR framework were nonetheless determined for the case studies by methods recommended by ProSafe, i.e. crystalline phase by XRD, surface hydrophobicity by contact angle measurement, surface chemistry by XPS, surface charge via zeta potential by electrophoretic analyses.

Tier 1_NEP Methods for “what is the NEP” properties

Description of “what is the NEP” was beyond scope of the ProSafe review but does not require methods other than the highly established elemental composition analysis (e.g. full digestion, then ICP-OES) and the morphological characterisation of a TEM cross-section to assess the dispersion state (examples in Figure SI_3, criteria in Table 2). Producers of NEPs would not need to measure the composition if it is known from the production process.

Tier 2 Methods for “where they go” properties

For the biodurability as central element of the „where they go“ assessment we follow the advice of ProSafe to use the OECD draft TG on “solubility in aqueous media”, but restricted here to the „screening method“, which we use in Tier 1 to assess the solubility in water (documented as mg/L value, assessed in the OECD draft TG as „% dissolved“) and in Tier 2 to assess the release of ions in a relevant medium (in mg/L metric, since essential for ecotoxicity grouping). Once the intended use of the NEP is known, the release and the affected environmental compartment can be identified and in case of aquatic ecotoxicity the dissolution in the medium for the test organism can be determined [24]. We agreed with ProSafe that the „dissolution rate in physiological fluids [needs] to be further developed“, and showed on 24 (nano)forms of 7 substances that the flow-cell method as described by ISO:TR19057:2017 correctly predicts materials with low human *in vivo* and *in vitro* biodurability, and differentiates between nanoforms.[53, 54] The detection of dissolved ions provides the ng/cm²/h metric recommended by Oberdörster *et al.*[55], is grouped by order of magnitude (i.e., decadic ranges), and the group descriptors include the assessment of transformation via the TEM analysis of remaining solids.[54] The „extended“ test of the draft TG was not deemed useful in comparison.

Regarding the transport as another aspect of “where they go” properties, nanoGRAVUR does not endorse the full scope of OECD TG 318:2017, which requires to perform 54 measurements for 1 property per each NF, but instead we restricted the test to the specifically relevant medium to capture the dispersion stability by homoagglomeration. It is anyway not obvious how the multidimensional stability

diagram over pH, NOM and Ca^{2+} variation would be reduced to a simple descriptor to quantify the similarity of different NFs.[35] Reviews have considered methods to rank NEP matrix materials by their release rates during professional handling or consumer use. Our choice of tensile strength is highly standardised, but only relevant for mechanical stresses. For other release concerns we recommend to adhere to the stepwise decision-making process by the ISO draft TR of PG29, which would e.g. recommend the interlaboratory-tested NanoRelease protocol to assess and compare the form and rate of release by environmental weathering.[56]

There are numerous methods for the determination of dustiness data.[57-59] The most common methods (with conventional mg/kg metric), which are standardised and also proposed by ECHA, are the rotating drum method (RD, EN 15051-2:2013) and the continuous drop-down method (CDD, EN 15051-3:2013). Since dustiness levels depend strongly on the used method,[60, 61], different methods were tested concerning suitability, including the CDD method of EN 15051-3:2013, the small rotating drum method ([57], CEN/TC 137; in #/kg metric), the fluidizer method especially designed for fibres [58, 59] (in #/cm³, #/kg or #/kg/s metric), and a dustiness equivalent method (DEM, in mg/kg metric) that is designed to mimic the CDD method by evaluating intrinsically measured size distributions (see SI chapter 2.4.1). Due to different assets and drawbacks of the methods and strongly limited comparability, no single method could be excluded or preferred, and it is recommended that for grouping purpose only data of one dustiness method should be used. For the case studies, results based on the DEM method were chosen, since DEM was able to provide dustiness data for most of the test substances in the metric proposed by ECHA (i.e., in mg/kg metric). For fibres, the average concentration was determined, in accordance with the metric of most exposure limit values.

Tier 2 Methods for "what they do" properties

To assess „what they do" via surface reactivity, we combine the two abiotic assays of electron spin resonance (ESR, also known as electron paramagnetic resonance, with the DMPO and CPH spin traps as described by ISO/TS-18827:2017) and the very sensitive Ferric Reduction Ability of Serum (FRAS) assay [62] with the cell-based NR8383 macrophage assay. The NR8383 assay uses four standardised read-outs (LDH, ROS, GLU, TNF) and was pre-validated against *in vivo* inhalation studies.[63] We also found strong correlation between the abiotic assays and cell-based protein carbonylation, but did not see a necessity to complicate the scheme by yet another redundant assay. (Bahl *et al*. in preparation) We slightly disagree with ProSafe on the strategy and terminology to assess surface reactivity. Redox potential and band gap are not synonyms as it seems to be suggested by Steinhäuser *et al.*[32] Instead, the LUMO band energy may be a more relevant parameter,[21] but even that remains a proxy for the actual biological oxidative damage. Hence, we do not define methods for any of the proxies, but group by the directly determined biological oxidative damage (sBOD).

For specific purposes, or when the evidence from Tier 2 remains inconclusive, the user is advised to overrule Tier 2 by a more specific testing in Tier 3. It is intended as backup with methods that are specific to a certain purpose of grouping, and that may be performed under conditions that are specific to a certain intended use with a scenario of emission or exposure. Human toxicity testing in Tier 3 may include *in vivo* studies, preferentially by OECD test guidelines, to support the similarity of different NFs and would overrule dissimilarity of screening methods in Tier 2. If exposure is relevant for the purpose of grouping, then Tier 3 may compare the similarity of personal exposure at specific workplaces. If environmental fate is central for the hypothesis of grouping, then Tier 3 might rely on environmental fate and transformation, for which OECD guidelines (e.g. TG312) exist but are not yet validated and standardised for nanomaterials. The Tier 3 is not fully detailed here, as the escalation to Tier 3 is anticipated to be the exception rather than the rule.

Two major method gaps remain: We believe that hetero-agglomeration can be a predictive parameter for toxicity to environmental species such as algae. Therefore a screening method was developed that indicates the attachment efficiency to algae[25]. However, laboratory comparison tests and standardisation are still missing. It is also believed that the rigidity of fibres is predictive of adverse pulmonary effects,[64] but all methods are exploratory. An ongoing BAuA research project tests an approach based on curvature analysis and oscillatory measurement and may be one of the possible implementations.

Although the nanoGRAVUR framework can serve multiple purposes, it remains consistent with literature for the purpose of grouping by occupational hazard of handling nanomaterial powders: Here the nanoGRAVUR framework and method selection coincide with the DF4nanogrouping, except that the Tier 2 dissolution is now assessed by an improved methodology of dissolution *rates* in a different metric (ng/cm²/h instead of mg/L).[19, 20] Thus, via the recommendation of the DF4nanogrouping method selection in the ECHA guidance,[16] also the nanoGRAVUR method selection should be a defendable implementation of the ECHA grouping guideline. In comparison, the general ECHA physico-chemical guidance R7.1, that is also recommended by the ECHA grouping guidance[16], allows many alternative methods. We concluded that similarity or dissimilarity can only be substantiated by identical methods, as supported also by the EPA guidance on the differentiation of "discrete forms".[1]

We note that the determination of extrinsic properties by „functional assays "is as such not an innovation. Oomen *et al*. remarked that for molecular chemicals the partition coefficient log k_{OW} is an example of a functional assay with very high regulatory acceptance, combining high predictivity and practical value.[8] We believe that for (nano)particles the assays for surface affinity, surface reactivity and dissolution rate will become just as important. Benchmark nanomaterials and benchmark nano-enabled products are essential to achieve reproducible groupings across different labs with slightly

differing equipment (e.g. for dustiness, sanding, dispersion stability, reactivity).

Table 2 Properties, harmonised methods of analysis, descriptors, and benchmark materials as proposed by the nanoGRAVUR grouping-framework for nanomaterials. Additional information on methods is provided in the SI. The optional scoring of the possible values of the descriptors implements a data reduction to indexed property bands.

Properties	Methods	Descriptors, [Metric]	Benchmark-Materials	Optional: Data reduction to indexed property bands
Primary particle shape	NanoDefine methodology (consistent with ECHA nanoforms)	minimum external dimension [nm] + aspect ratio [unitless]	NanoDefine IRMM-repository	sphere 1, rod 2, platelet 3, fibre 4
Primary particle dimension			NanoDefine IRMM-repository	spherical and ∅<10nm 1; other 2
Rigidity (for fibres)	No valid method established yet	modulus of elasticity [MPa] (for MWCNT: diameter [nm])	NM400 (10 nm, non-rigid) // NM401 (67 nm, rigid)	not established
GHS CLP (bulk) humantoxicity, ecotoxicity	identify composition by XRF (or ICPMS, XRD), compare to CLP of Bulk if existent.	Consider impurities > 1%	Not required	
Solubility in water	OECD TG draft: in 5 mM $NaHCO_3$, pH7 at 10mg/L, 24h	Document mg/L of metal ion; The TG assesses % dissolved	TG specifies CuO (<50nm, SA: 29 m^2/g, non-coated). On CuO (PlasmaChem): 0.32mg/L dissolved	0 not significant / 1 low (<1mg/L) / 2 mid (<10mg/L) / 3 high (<50mg/L)
Physico-chemical hazards (bulk)	Bulk GHS CLP	H-phrases	Not required	
Product classes and application scenarios	None (descriptive)		Sunscreen // plastics (solid polymers) // cement	consumption by consumer // contained in consumer products // industrial or professional use
Specific application in NEP: State of dispersion of ENM	Assignment to three fixed categories, which determine the disperse system as well as the type of embedding and agglomeration		Not required	Disperse system: 1 composites // 2 suspensions // 3 powder Embedding into a matrix: 1 complete embedding // 2 partly embedding // 3 attachment // 4 isolated Agglomeration in matrix: 1 highly agglomerated // 2 slightly agglomerated // 3 individualised
Specific application in NEP: content (g/g) of ENM in NEP		Mass-% ENM in NEP	Not required	0 low/ 1 mid/ 2 high with low<5wt%; mid: 5 to 50 wt%; high>50wt%
Dustiness	EN 15051 Methods (RD, CDD) and alternative methods with similar strain intensity (e.g. SRD, SHA, DEM) or FLU (esp. for fibres)	dustiness coefficient dependent on mass and number [mg/kg, #/kg] factor of emission in number metric	Fibre benchmarks in FLU method: NM400 (low dusting tendency: 150/mg/h), NM401 (high dusting tendency: 8000/mg/h)	(0 low/ 1 mid/ 2 high) Particle ranges in DEM method: Inhalable: low < 4000 mg/kg, moderate 4000-15000 mg/kg, high >15000 mg/kg Thoracic: low <1847 mg/kg, moderate 1847-5000 mg/kg, high >5000 mg/kg Alveolar: low <70 mg/kg, moderate 70 mg/kg – 300 mg/kg, high > 300 mg/kg
Agglomeration of ENM upon application of NEP	For fibres: dustiness For sprays: intended application. Each with morphological analysis	fibres: volumetric share of constituent fibres in agglomerates in relation to amount of single fibres within the dust (%)	Fibre benchmarks: ARIGM001 ($X_V = 0.003$), NTX-3 ($X_V = 0.913$)	Fibre ranges < 0.04 (low), 0.04 - 0.4 (medium), > 0.4 (high)

	of the dust			
Resilience of NEP matrix	For mechanical stress: tensile elongation (ISO method)	tensile strength [MPa] or elongation at break [%]	tensile strength (inverse to sanding release rates): • Low resilience (< 10 N/mm²): cement • mid resilience (10 – 100 N/mm²): Epoxy, PA, Acrylic • high resilience (> 100 N/mm²): steel, aluminium alloy	
Critical dimensions upon exposure	Dustiness + SEM analysis of aerosol sample: form	Amount of WHO-fiber-like objects from total number (%)	NM400: 0.4% NM401: 20.4%	Not established
Ion-releasing	OECD TG draft: screening method 10 mg/L ENM in relevant medium	dissolved ions [mg/l]	ion releasing if >0.1 mg/L metal ions. CuO in env. medium >0.8 mg/L Cu^{2+}	0: no / 1: yes
Dissolution in relevant media	flow-through dissolution: ICPMS quantification of ions	Rate k [ng/cm²/h]	For lysosomal dissolution: • Non-persistent, high dissol. (ZnO NM110) • Non-persistent, significant transformation ($BaSO_4$ NM220) • Low dissolution, significant transformation (SiO_2 NM203) • Low dissol., low transform. (CeO_2 NM212) For environmental transformation: $BaSO_4$ NM220 (particles dissolve but become more crystalline)	For lysosomal dissolution: 1: non persistent high dissolution k>100 ng/cm²/h 2: non persistent significant transformation k=1-100 ng/cm²/h 3: low dissolution significant transformation k<1 ng/cm²/h 4: low dissolution, low transformation k<1 ng/cm²/h For enviromental transformation : 0: non persistent, high transformation 1: no / low transformation
Transformation "changes of what they are"	flow-through dissolution: TEM, (SAD, XPS) detection on remaining solids	Comparison of shape and size (optionally also crystallinity) before/after dissolution testing		
Homo-agglomeration	TG318 in relevant medium (instead of 3*3 Ca*NOM media)		TiO_2 NM105 10-90%, Ag NM300 >90% (from TG318)	<10% unstable (0) / intermediate (1) / >90% stable (2) (from TG318)
Affinity (hetero-agglomeration)	tbd. possibly Geitner *et al.* ES&T 2016 or microscopic qualitative analysis relative to benchmark	attachment efficiency (α) or fraction of attached particles	tbd.	no 0 / yes 1
Mobility	soil columns based on OECD TG 312	transport distance [% of total length]	tbd	1:high or breakthrough (100%), 2: mobile in soil column, 3: no mobility
Reactivity (abiotic)	ESR (ecotoxicity), ESR+FRAS (human toxicity)	ESR: relative to LoD if 1.3*negative control: mBOD(CPH) and mBOD(DMPO) FRAS: relative to LoD and positive benchmark: sBOD and mBOD	ESR: $BaSO_4$ (neg) vs $CuSO_4$ (pos) FRAS positive benchmark: Mn_2O_3 (sBOD=2866 nmolTEU/m² ENM).	additive scoring of ESR/FRAS: significant above LoD (+1/3 point), above 10% of positive control (+1 point) 1: low, 2: mid-low, 3: mid-high, 4: high
Reactivity (*in vitro*)	NR8383 dose response NRK-52E protein carbonylation	• LDH (LOAEC) [µg/mL] • GLU (LOAEC) [µg/mL] • TNF (LOAEC) [µg/mL] • H2O2 (LOEAC) [µg/mL] Carbonyl intensity rel. to positive control.	No benchmark required CuO (positive control)	Score = number of vectors with significant effects at surface dose of 0.006 m²/mL Not established

Additional properties that were tested in case studies but not selected for the nanoGRAVUR framework				
Surface Chemistry	Descriptive	Not required	Not required	1 untreated; 2 hydrophilic functionalization; 3 hydrophobic funcationalization; 4 core-shell coating
Hydrophobicity	Water sessile drop	contact angle (°)	Not required	H'il = hydrophilic, 0 to 90°; H'ob = hydrophobic, >90°
Surface charge	Electrophoretic mobility	Zeta potential at pH7, mV	Not established	Not established
Reactivity (photo-)	Rhodamine-B-degradation	Photon efficiency (unitless, %)	Not established	Not established

Table 3 Numerical results of the case studies and OECD benchmark materials for both the ECHA recommended properties and the additional nanoGRAVUR harmonised properties. For each property, results were obtained for all materials by the same method (Table 2). For the properties that were selected for the nanoGRAVUR grouping framework (Table 1), the numerical values are assigned to bands in Table 4. Further measurements (e.g. transformation by TEM scans) provide images that are evaluated directly towards bands as shown in Table 4. "n.s." indicates that the result was not significant against the limit of detection. White cells are data gaps; the goal was to test grouping hypotheses, and not to fill all data gaps. Grey cells indicate that the method is not applicable. For each property and descriptor, the results are color-coded between the negative and positive benchmark materials (if defined in Table 2), and otherwise between the minimum and maximum values. For surface reactivity (*in vitro*), light color without numerical value documents that the property was determined, but that this descriptor was not significant up to a dose of 180 µg/mL.

descriptors		DPP_nano	DPP_non-nano	DPP_premixed	CuPhthalo_nano	CuPhthalo_halogen	Fe2O3_nano_A	Fe2O3_nano_B	Fe2O3_larger	SiO2_untreated	SiO2_amino	SiO2_phosphonate	NM203_SiO2_h'phil	GBS	GGBS	HKP_GGBS_hr	CEM	NM211_CeO2	NM212_CeO2	NM110_ZnO	NM111_ZnO	CuO	NM105_TiO2_nano	NM102 - TiO2_active	NM104 - TiO2-coat.	TiO2_non-nano	NM400_CNT	IRMM382_CNT	Carbon Black	Graphene_monolayer	Graphen_multilayer	NM600_Bentonit	WS2	IRMM385_Kaolin	Kaolin	IRMM381_BaSO4 non-nano	IRMM387_NM220_BaSO4	Mn2O3
Composition	CAS no.		84632-65-5		147-14-8	132-53-5	1309-37-1			7631-86-9				65996-69-2			65997-15-1, 68475-76-3	1306-38-3		1314-13-2		1317-38-0	1317-80-2	1317-70-0	13463-67-7		308068-56-6		1333-86-4	1034343-98-0		1302-78-9	12138-09-9	1332-58-7	1332-58-7	7727-43-7	7727-43-7	1317-34-6
Primary particle dimension	median diameter, nm	43	233	400	17	39	12	37	48	15	15	15	26.2	206	833	1234	399	15	40	42	80	24	21	34.8	29.6	204	15.6	10				1362	100	181	279	234	32	36
Specific surface area (BET/VSSA)	BET, m^2/g	94	16	17	53	69	107	30	12	200	200	200	213	1.1	2	2	1.8	66	27	12	14	34	51	80	60	15	254	234	56.5	559	17.6	52		16	24	2.5	41	20
Surface Chemistry (measured)	% C	77.1	79.4	73.5	80.5	55.6	15.7	50.7	23.9	4		5						14.1	79.9	30	67.9	7	9	23.4	16.3		99										17	
	% O	10.9	19	9.5	9	0.8	54.2	33.7	49.6	66		66						57.2	17.7	38	24.3	47	63	50.7	63.5		1										52	
	% N	5.9	5.1	8.1	8.5	11.8																															1	
	% metals		0.3	0.4	1.4	1.4	28.3	15.6	24.6	30		29						28.7	2.4	38	4.3	46	29	18.6	20.2		<1										13	
	% non metals	6.1	5.2	8.6	0.7	30.3	1.8													3	3.5																17	
Surface charge (zeta-potential)	Zeta potential at pH7, mV	-16	-41	-30.4	-11	-38	-27	18	-55	-39	0	-43	-24					-24	15	30	-25	-34	-17	-33	26	36			-16.5	-45.4	-40.7	-31	-49.2		-53	-37	-37	-5
Hydrophobicity	water contact angle	135	136	103	138	163	10	10	10	0	0	0	44	10	10	10	10	10	60	10	152	10	60	10		10	140		148	93	79	10		10	10	10	10	
NEP class & intended use scenarios	lifecycle release tested	automotive coating											cement / concrete							sun screen			conc rete		sun scre.	coat ing	polymer for mobility						polymer for mobility				
Specific NEP: g/g content of ENM	% g/g	3	3	3	3	3	3	3	3				2	10	10	10	10			10	10		2		10	5	1	1	1	1	1		1		25			
Dustiness	g/kg inhalable	23	9	23	10	4	30	8	4				146	1	1	1	2					1	12			4			9		43				7			
	mg/kg thoracic	3064	428	1498	348	120	5273	358	155				19907	11	27	10	25					13	554			194			885		3867				497			
	mg/kg respirable	51	0.6	6.6	0.4	0.1	220	0.4	0.1				1176	0.00	0.01	0.00	0.01					0.00	1.1			0.1			3.6		40				1			
	mean #/cm^3 (for fibres)																										179											
Critical shapes upon exposure (for fibres)	SEM of aerosol																										0.9											
Agglo. of ENM upon NEP application	SEM of aerosol																										0.033											
Dispersion stability (environ. homoaggl.)	% stable after 6 hours in ADaM				7	10	17	2	14	74	95	82	48					9	25	29	36	9	11	5								22			15		53	
Solubility in water	OECD screening LoD or value, metal ion, ppm				ns	ns	ns	ns	ns	27	13	12	56					0.1	0.1	3.6	3.3	97	ns	ns	ns	ns						0.1			0.8		6	0.2
Ion releasing in relevant environ. media	OECD screening LoD or value, metal ion, ppm						ns	ns	ns	51	27	55						ns	ns	2.3	2.4	0.8																
Dissolution rate in relevant human media	lysosomal dissolution rate k [ng/cm2/h]				0.76	0.42	0.04	0.04	0.1	0.2	0.27	0.45	0.4					0.14	0.06	204	177	283	0.013	0	0.013	0						0.65			1.3	53	10	2
Mobility (in soils)	% column transported				10		90	90										100	10													100					100	
Reactivity (abiotic)	ESR_CPH: mBOD x-fold of D2O control	0.8	0.8	0.9	1.2	1	0.5	0.8	13.9	4	1	2.2	2.2		1	0.7	1	10.6	7.9	3.4		39.7	2.3	2.7	1.6	0.9	55		3.6	150	3.1	2.8	0.7		1.6		2	15.8
	ESR_DMPO: mBOD x-fold of D2O control	0.8	0.8	0.8	1.6	7	1.1	0.8	4.3	11	21	19	19		1.3	1	0.9	1.3	1.4	2.8		14.1		0.9	0.9	1	1.1		0.9	0.75	1.15	1.3	1.2		1		2	2.3
	FRAS mBOD [nmolTEU/mg ENM]	2	1	2	3	2	5	1	1	3	4	8	4					4	6	43	25	326	5	13	0.7	2	23		8	110	14	6	68	0.2	4	0	0	78
	FRAS sBOD [nmolTEU/m^2ENM]	3.4	2.4	0	12	18	44	15	34	14	18.9	8						14	8	151	20	9486	19	6.3	12	14						87		13	16	0	7	2866
Reactivity (in vitro) NR8383 cells	LDH (LOAEC) [µg/mL]		90	180	90	180	90	180		22.2	45	90	90	180				90	90	2.8	5.6	2.8	90	180	90					45	11.3	90	90	45			45	
Reactivity (in vitro) NR8383 cells	GLU (LOAEC) [µg/mL]		180		90	180		180		45	45	180	22.5	180	180	180	180	180	180	11.3	11.3	2.8	90	180	90		90			180	45	22.5	45	90	45	180		180	
Reactivity (in vitro) NR8383 cells	TNF (LOAEC) [µg/mL]						180			22.5	22.5	90	180					22.5	22.5	11.3	22.5		90	180	90					22.5	45				90			
Reactivity (in vitro) NR8383 cells	H2O2 (LOEAC) [µg/mL]							45		45	180	45		90	45	22.5			22.5			22.5									180	5.63	22.5	180	45			22.5
Reactivity (photo-)	photon efficiency [%]	0.49	1.1	0.82															1.97	16		10															

Table 4 Data reduction of the case studies and OECD benchmark materials: each NF is assigned to a band, represented by a score 0 to 4. The scoring system is specified in Table 2. The similarity between different NFs of the same substance is assessed by the scores of those properties that are relevant for the grouping hypothesis, and as selected by Table 1 for the different purposes of grouping.

	DPP_nano	DPP_non-nano	DPP_pernixed	CuPhtalo_nano	CuPhtalo_halogen	Re2O3_nano_A	Re2O3_nano_B	Re2O3_bigger	SiO2_untreated	SiO2_amino	SiO2_phosphonate	NM203_SiO2_hydrphil	GBS	GGBS	HKP_GGBS_hr	CEM	NM211_CeO2	NM212_CeO2	NM110_ZnO	NM111_ZnO	CuO	NM105_TiO2_nano	NM102 - TiO2_active	NM104 - TiO2-coat	TiO2_non-nano	NM400_CNT	IRMM832_CNF	Carbon Black	Graphen einlagig	Graphen_mehrlagig	NM600_Bentonit	WS2	IRMM885_Kaolin	Kaolin	IRMM881_BaSO4 non-nano	IRMM887_NM220_BaSO4	Mn2O3
Primary particle shape	1	1	1	1	1	2	1	1	1	1	1	1	1	3	3	3	1	1	1	1	1	1	1	1	1	4	4	1	3	3	3	4	3	3	1	1	1
Primary particle dimension	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2	2	2	2	2	2	2	2	2	2
Composition	-	-	-	-	-	315-319-335	315-319-335	315-319-335	-	-	-	-				302-315-318-335	-	-	410	410	302-410	-	-	-	-	319-335		351	319-335	319-335	330	315-319-335	-	-	-	-	302-312-315-319-335
Surface Chemistry (descriptive)	1	1	2	1	1	1	1	1	1	2	2	1	1	1	1	1	1	1	1	3	1	1	1	3	4	1		1	1	1	1	1	1	1	1	1	1
Hydrophobicity	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0		0	1		1	1	0	0			0		0	
NEP class & intended use scenarios	automotive coating											cement / concrete							sunscreen			concrete		sun scre.	coating	polymer for mobility						polymer for mobility					
Specific NEP: dispersion state of ENM	1	1	1	1	1	1	1	1				1	1	1	1	1			2	2	2			2	1	1	1	1	1			1		1			
	1	1	1	1	1	1	1	1				1	1	1	1	1			1	1	3			1	1	1	1	1	1			1		1			
	3	3	3	3	3	3	3	3				2	2	2	2	2			2	2	2			2	3	3	3	3	3			3		3			
Specific NEP: g/g content of ENM	0	0	0	0	0	0	0	0				0	1	1	1	1			1	1		0		1		0		0	0	0		0		1			
Dustiness	2	1	2	1	0	2	1	0				2	0	0	0	0					0	1			1	0		1		2				1			
	1	0	0	0	0	2	0	0				2	0	0	0	0					0	0			0			0		1				0			
	0	0	0	0	0	1	0	0				2	0	0	0	0					0	0			0			0		0				0			
Dispersion stability (environ. homoaggl.)				1	0	1	1	1	2	2	1	1					1	1	1	1	1	0		0										1		1	
Mobility (in soils)				3		2	2										1	3				3										1				1	
Attachment HBS:APBSto algae				0	0	1	1	0	0	0	0						1	1				1		1													
Solubility in water				0	0	0	0	0	2	2	2	3					0	0	2	1	3	0	0	0	0							0		1		2	
Ion releasing in relevant environ. media				1	1	0	0	0	1	1	1	1					0	0	1	1	1	0	0	0	0	0	0	0	0	0	1	0		1		1	1
Dissolution (environmental perspective)				1		0	0	0	0	0	0	0					1	1	0	0	0	1	1	1								1		0		0	0
Transformation (environmental perspective)				1		0	0	0	0	0	0						1	1	0	0	0	1	1	1	1							1				1	
Dissolution & Transformation (human perspective)				3	3	4	4	4	3	3	3	3					4	4	1	1	1	4	4	4	4						3			3	2	2	
Reactivity (abiotic)	0	0	0	1	1	1	0	2	2	2	2	2					1	1	4	2	4	1	1	0	0						2		1	1	0	1	3
Reactivity (in-vitro), human perspective	0	2	1	2	0	0	3	0	2	1	0	2	3	2	2	1	2	2	3	3	3	3	0	3	0	0		0	0	0	2	0	3	4	1	0	2

Proof of concept (Case Studies, Banding and Calibration)

We applied the framework to 34 (nano)forms of 17 substances. Table 3 shows the numerical values of the material properties (Tier 1, Tier 1_NEP and Tier 2 from Figure 1), supplemented by representative SEM or TEM scans in Figure SI_1. The case studies cover particle, platelets and fibrous shapes, with sizes from 10 nm to >1µm, BET surface areas from 1 to >500 m²/g, surface chemistry: carbon from 4 to 99%, surface charge from -55 to +36 mV, and hydrophobicity determined as water contact angles from <10° to 163°. This lends us to believe that the case studies explore a relevant portion of the NF design space.

Boundaries for the numerical ranges that define groups depend on the purpose (perspective) of grouping and may be universal or substance-specific. The nanoGRAVUR framework does not yet support a quantitative measure of similarity, nor criteria for sufficient similarity to substantiate read-across. But our diverse case studies allow an analysis how criteria of similarity impact the groups that result: One could consider the similarity between different NFs of the same substance as "ideal" if the different NFs are assigned to the same bands for *all* Tier 2 properties that are selected for the grouping purpose by Table 1. In a more pragmatic weight-of-evidence approach, one would jointly assess several properties that are relevant for the grouping purpose; such approach would implement the stepwise ECHA process of hypothesis formulation, data gathering and hypothesis substantiation for regulatory grouping.[16] In any case, Tier 3 can be used to overrule Tier 2, as exemplified in case studies.

In the following, we assess the similarity within substance families with the properties and descriptors given in Table 1 and Figure 1 for the occupational and environmental grouping perspectives. We compare against *in vivo* and OECD guideline studies (Table 5) of human and environmental hazards, by well-established inhalation and relevant aquatic and soil organisms. We then perform via the rules given in Table 2 the data reduction to simple descriptors and order-of-magnitude (decadic) bands, resulting in Table 4, and again assess the similarity – this procedure calibrates the banding and grouping against regulatory testing. Below, we discuss which properties are very sensitive to different NFs, and how this might impact the conclusions on grouping of such NFs.

$BaSO_4$: The NF and the non-nano-form of $BaSO_4$ share the same dissolution rate, and are similarly low reactive, both under abiotic and *in vitro* conditions. The NR8383 assay ranked the non-nano-form as being more reactive, but still not as "active". The Tier 2 human perspective approach (Figure 1) would thus recommend a common group for both forms. The *in vivo* STIS of NM220 $BaSO_4$ confirms low hazards at aerosol concentrations as high as 50 mg/m³ (Table 5). The ECHA guideline, which requires a read-across process and, moreover, does not recommend a common group for nano- and non-nano-forms, appears to be overly conservative in this case study.

Organic pigments: The three DPP pigments form are another substance family with low abiotic reactivity. However, the non-nano-form showed two positive results in the macrophage assay (LOAECs below the surface area-based threshold of < 6000 mm²/mL), and was considered to be „active".[63] The non-nano-form was, therefore, a candidate material for Tier 3 *in vivo* testing. However, at that stage of analysis a short-term inhalation study (STIS) had already shown that both forms of DPP elicit no adverse effects up to 30 mg/m³,[68] (Table 5). Due to this similarity a joint assessment (grouping or read-across) of nano and non-nano DPP pigments appears justified. [20]

Table 5 *In vivo* human toxicity and/or ecotoxicity testing on the case study ENM and benchmark materials. The existence of the data in this table was initially the reason for selecting the case studies and provided the basis to evaluate and discuss the validity of the framework. Inhalation hazards tested by short-term inhalation screening on rats (STIS, 5d exposure at aerosol concentration indicated, 21d recovery); results indicated the NOAEC or its lower limit by the highest dose tested without adverse effects. [20, 68-74]Environmental hazards tested by Algae (OECD 201; Raphidocelis subcapitata), Daphnia magna (OECD 202), Zebrafish embryo (FET, OECD 236) and soil microflora (ISO 15655), with results given as EC50 or highest dose tested.[24, 25] n.d. = not determined.

	NOAEC (rat, STIS) mg/m³	EC50 (algae) mg/L	EC50 (daphnids) mg/L	EC50 (FET) mg/L	EC50 (soil micro flora) mg/L
DPP_nano	>30	n.d.	n.d.	n.d.	n.d.
DPP_non-nano	>30	n.d.	n.d.	n.d.	n.d.
CuPhthalo_nano	>30	>100	>100	>100	>1000
CuPhthalo_halogen		>100	>100	>100	>1000
Fe2O3_nano_A	30	3.6	>100	>100	>1000
Fe2O3_nano_B		2.4	>100	>100	>1000
Fe2O3_larger	30	111	>100	>100	>1000
SiO2_untreated	2.5	14	>100	>100	>1000
SiO2_amino	>50	29	>100	>100	>1000
SiO2_phosphonate	>50	46	>100	>100	>1000
NM203_SiO2_hydrophil	1.0	n.d.	n.d.	n.d.	n.d.
NM211_CeO2	<0.5	8.5	>100	>100	>1000
NM212_CeO2	<0.5	5.6	>100	>100	>1000
CuO (PlasmaChem)	0.6	1.4	0.3	≈30% effect at 100 mg/L	~1000
NM110_ZnO	n.d.	0.1	3.4	>100	118
NM111_ZnO coated	0.5	0.1	8.3	>100	173
NM105_TiO2_nano	<2	4.7	n.d.	n.d.	n.d.
NM104	n.d.	63	n.d.	n.d.	n.d.
NM400_CNT	<0.5	n.d.	n.d.	n.d.	n.d.
NM220_BaSO4	50	n.d	n.d	n.d	n.d
Quartz DQ12	0.1	n.d	n.d	n.d	n.d

TiO_2: Among the intrinsic properties of the four tested TiO_2 compounds we observed differences not only in particle size (21-204 nm) and surface area (15-80 m²/g) but also in surface coating. Here the UV-active TiO_2 NM102 and NM105 showed a slightly higher reactivity in the abiotic test (band 1) than the NM104_coated and the non-nano-form. The *in vitro* assay dose -response is actually not very different between the three NFs (Table 3), but the scoring via specific surface area assigns band 3 (NM105, NM104) and band 0 (NM102), respectively, and is thus suggestive of dissimilarity in contrast to the similar abiotic reactivity in surface metric (FRAS sBOD, Table 3). The non-nano-form, which has an alumina coating, was similar to the NM104 coated NF in the absolute values of reactivity, and had no significant *in vitro* reactivity. The other extrinsic properties of TiO_2 materials were highly similar, with a

dispersion stability <10% after 24 h, and, furthermore, neither dissolution and nor transformation. As shown elsewhere [25], the attachment efficiency varies significantly between different TiO_2 NFs. Moreover, in the NR8383 macrophage assay TiO_2 nanomaterials elicited divergent effects. In summary, neither the occupational nor the environmental perspective indicated enough similarity of the different TiO_2 NFs. This finding was attributed to the different crystallinities and coatings and does not rule out that more homogeneous selections of TiO_2 NFs may be sufficiently similar to justify groupings.

SiO_2: In the Tier 2 environmental perspective, the different SiO_2 NFs are similar to each other in the relevant property bands of mobility in soils, dispersion stability, abiotic reactivity, and attachment to algae (Table 3). Because the dispersion stability ranges just around to 90 % group cut-off suggested by TG318, the SiO2_amino with 94 % stability is not in the same band as the SiO2_untreated with 74 % or the NM203 with 48 %, respectively. This does not need to prevent a grouping but would necessitate both aquatic and sediment testing in Tier 3 ecotoxicity. Hund-Rinke *et al*. previously published the ecotox-scheme for the grouping of NMs, which was based on the properties "Ecotoxicity of bulk material", "Ion release", "Reactivity" and "Morphology/Size". The EC50 results of aquatic species confirmed the similarity of SiO2_untreated and SiO2_amino, where algae as most sensitive species had EC50 of 14.1 and 29.2 mg/L respectively.[24] In the Tier 2 human (occupational) perspective, the different SiO_2 NFs all share a slow (<1 ng/cm²/h) dissolution and significant transformation by gel formation.[54] All are assigned to the same dissolution/transformation band (Table 4). However, their reactivity differs: the SiO2_untreated has two LOAEC in the NR8383 assay below the surface area-based threshold of < 6000 mm^2/mL. It is thus assigned two scores for reactivity (*in vitro*) and would be considered as „active" in the DF4nanogrouping, similarly to NM203. Other colloidal NFs have fewer LOAEC below that threshold, and are considered as "passive" in the DF4nanogrouping.[63] The abiotic reactivity is quite heterogeneous between the FRAS and EPR parameters in the numerical values (Table 3) but averages out to a combined band 2 for all tested silica NFs (Table 4). Overall, the differences in *in vitro* reactivity would prevent grouping in Tier 2 and would necessitate Tier 3 testing. *In vivo* studies (by STIS) confirmed that the NFs are different (Table 5) and would indicate SiO2_untreated would be a suitable source for the SiO2_amino and SiO2_phosphonate as target NFs.[75, 76]

Aluminosilicates: Comparing the SiO_2 (nano)forms with the aluminosilicates (kaolin, bentonite), we found similarities in dissolution and transformation from both the environmental and human perspective, but differences in abiotic or *in vitro* reactivity (Table 3). Thus, platelet shaped aluminosilicate particles were more reactive in the *in vitro* reactivity (bands 2-4, Table 4) than their round shaped SiO_2 counterparts (bands 0-2), and even showed a higher cytotoxicity than our positive control Mn_2O_3. In contrast, the abiotic reactivity was lower for the aluminosilicates than for the SiO_2. Comparing the two Kaolin NFs, IRMM385_Kaolin (with BET of 16m²/g) has lower abiotic and *in vitro* reactivity *per BET* than the other Kaolin (with BET of 24 m²/g), but is still similar. Both Kaolins are significantly less reactive than the Bentonite in the NR8383 dose response (Table 3), but due to the high BET of Bentonite, the *in vitro* scoring system (compare SI) only results in a band 2 for Bentonite (Table 4). At present the high *in vitro* reactivity especially of bentonite is an unresolved issue. First *in vivo* studies show a high inflammatory potential of bentonite inside the rat lung (manuscript in preparation), thus confirming the NR8383 testing results. Possibly the layered structure of bentonite, rather than the platelet structure, interferes with the micro-milieu inside the phagolysosomal compartments, e.g. due to swelling, ion binding and/or osmotic challenges. However, the damage inferred by bentonite to the lung was transient, suggesting that the layered structure of bentonite transforms into a far less bioactive particulate. Understanding of the transformation processes *in vivo* may help. Studies on graphenes before/after reduction (thus comparing changes to composition at same shape) confirm that the chemical speciation is important to describe „where they go".[77] Clearly, nanoforms that share the shape of thin platelets can be very dissimilar.

Cu-compounds: We also compared three copper containing materials with similar particle size but different CAS numbers (Tier 1, Figure 1). CuO and two CuPhthalocyanines differed already with respect to their Tier 1 properties, e.g. by different elemental compositions (5-10 % Cu for the pigments and 46% for CuO) and by the GHS classification of bulk CuO as compared to no GHS classification of CuPhthalocyanines. Furthermore, CuO is hydrophilic whereas the other two materials are hydrophobic with contact angles up to 163°. The sizes of the primary particles of the three materials did not vary much with 17 and 39 nm for the CuPhthalocyanines and 24 nm for CuO. Within CuPhthalocyanine the Cu is strongly bound to the complex, making it almost insoluble with dissolution rates around 0.5 ng/cm²/h for both CuPhthalocyanines and 282 ng/cm²/h for CuO.[54] Accordingly, CuPhthalocyanines showed only limited reactivity within the FRAS, ESR and NR8383 assay with bands of 1 in the abiotic assay and between 0-2 in the *in vitro* assay. Whereas CuO quickly dissolved, setting free Cu-ions. CuO was the material with the highest reactivity in the abiotic test and a band of 3 in the *in vitro* test. It furthermore dissolved the quickest, even quicker than the ZnO materials at lysosomal pH. The drastic differences are confirmed by STIS, showing complete dissolution of CuO with NOAEC at 0.6 mg/m³ [73] vs. no adverse effects for CuPhthalo_nano up to the highest dose tested of 30 mg/m³ [20] (Table 5). The CuPhthalo_nano gave most likely false positive result in the NR8383 assay, since STIS showed a low hazard potential,[20]. Also, specific tests of ecotoxicity confirm the dramatic difference between Cu with an EC50 below 0.1 mg/L for algae and daphnids, vs. a non-detectable EC50 above 100 mg/L for the CuPhthalocyanines (Table 5).[24, 25]

Fe_2O_3 : For the three tested Fe_2O_3 compounds, different values for the intrinsic properties were measured, such as a range in BET from 12 to 107 m²/g, and furthermore the Zeta potential ranging from -18 to -55 mV for Fe_2O_3_nano_b and

Fe_2O_3_larger respectively (Table 3). Despite the significant differences in intrinsic properties, almost all extrinsic properties were found in the same bands in *Table 4*. E.g., all Fe_2O_3 (nano)forms were similar in low dispersion stability <10% after 24h. They also shared a very low dissolution below 0.1 ng/cm²/h in lysosomal conditions, without transformation.[54] However their reactivity differed: the abiotic reactivity band ranged from mid-low (2) to mid-high (3), but remained below 10% the positive control for all forms, hence they were "passive" in the DF4nanogrouping terminology.[19] In the NR8383 assay, Fe_2O_3_nano_B was the only form with significant *in vitro* reactivity, whereas the other two materials were not reactive. Due to this disparity, Tier 3 comparison was applied and STIS indicated a similar outcome for the NF and non-nano-form of Fe_2O_3 with respect to treatment-related microscopic findings after days 5 and 26 (Table 5).[68] Thus, grouping appears justified in the occupational perspective. The similarity of Fe_2O_3 materials was even higher for the environmental fate descriptors, which fall into the same bands for low dissolution, intermediate mobility in soils, intermediate dispersion stability, lacking transformation and ion release. Thus, the environmental fate of Fe_2O_3 materials appears very similar. Spatial proximity (e.g. by attachment) is required for reactivity-induced ecotoxicity.[33] The environmental hazard was estimated to be similar for the two NFs, but the attachment to algae is more pronounced for the non-nano-form.[25]. The EC50 results of aquatic species confirmed the similarity of Fe2O3_nano_A and Fe2O3_nano_B, where algae, as most sensitive species, had EC50 of 23.4 and 17.7 mg/L (*Desmodesmus subspicatus*)[24, 25] and of 3.6 and 2.4 mg/L (*Raphidocelis subcapitata*)[25] respectively. In contrast, the EC50 for Fe2O3_larger is extrapolated to be 111 mg/L (*Raphidocelis subcapitata*)[25] (Table 5). Grouping the Fe_2O_3 NFs for the purpose of demonstrating environmental safety is thus justified, but a discrepancy remains between the NFs and the non-nano-form.

CeO_2: The assessment of the properties of CeO_2_NM211 and CeO_2_NM212 showed that they had hardly any intrinsic properties in common, such that particle sizes (15 and 40 nm), BET (66 and 27 m²/g), contact angle (10° and 60°), and surface charge (-24 and 15 mV) were all different. Nevertheless, CeO_2_NM211 and CeO_2_NM212 behaved similar in many of the Tier 2 properties: Both NF had a rather low dispersion stability with <10% stable particles after 24 h, no transformation (band 1), low dissolution (band 1). It was furthermore found that both materials share band 2 in the *in vitro* reactivity assay, and band 1 for the combined abiotic assays. The NF of smaller size, NM211, had higher reactivity FRAS sBOD per surface. This small disparity in Tier 2 is an interesting case for the calibration of "sufficient similarity", as their similarity from the occupational perspective was confirmed by STIS in Tier 3 (Table 5), which furthermore revealed similar neutrophil influx at aerosol concentrations of 5 mg/m².[71] Results of the two Ceria materials were even closer related if the surface area of the particles was used as a dose metrics for inflammatory responses.[71] However, their environmental transport in soils was significantly different and prevents grouping for all environmental endpoints. The strongly reduced mobility in soils of NM212 may be attributed to the hydrophobicity and positive charge of this material. On the other hand, properties relevant for environmental hazard (release of ions, reactivity / attachment, shape) were similar, indicated by EC50 values to algae which amounted to 8.5 and 5.6 mg/L respectively (Table 5).[2425] Thus, grouping the two (or more) NFs of CeO_2 might not be justified, but read-across for specific endpoints, as envisioned by ECHA,[16] is justifiable.

ZnO: The metal oxide particles ZnO_NM110 and ZnO_NM111 were largely different from CuO with respect to chemical composition and intrinsic properties (primary particle size 42-80 nm for both ZnO and 24 nm for CuO and BET 12-14 m²/g and 34 m²/g respectively). Nevertheless, all three materials shared bands with <10% dispersion stability, and the highest solubility, and were non-persistent/high transformation materials with high dissolution rates. STIS of ZnO NM111 [72] and CuO [73] confirmed the high solubility *in vivo* and indicated a fast clearance from the lung. Furthermore, ZnO und CuO exhibited a high abiotic (band 4) and *in vitro* reactivity (band 3). Although grouping across substances is not permitted by ECHA for regulatory purposes,[16] an industrial risk estimation based on these results would consider ZnO und CuO similar and would suggest the same measures for worker protection at production sites.[72, 73] They are also the only materials with significant photon efficiency in the Rhodamine-based assay on photo-reactivity, which is proposed by the ECHA grouping but not incorporated in the nanoGRAVUR grouping scheme.[16] A large similarity of ZnO NM110 und ZnO NM111 was also suggested by the EC50 values for algae (0.1 mg/L) and daphnids (3.4 to 8.3 mg/L) (Table 5).[24, 25] Within a substance, read-across from a soluble Zn salt to the ZnO NFs seems justified due to the biological similarity and high (>100ng/cm²/h) dissolution rates.

NEPs: The nanoGRAVUR scheme also offers a comparative risk estimation of NFs which were integrated in nano-enabled products (NEPs) for automotive coatings, clinker-reduced cements, cosmetic sunscreen and lightweight polymers. Representative photographs and cross-section SEM or TEM analyses of the NEPs are shown in Figure SI_2 and SI_3. The values of "matrix resilience" are not reported in Table 3 because they were not re-measured for each NEP, and because some NM were integrated in several different matrices, for release testing. As expected, incorporation into solid matrices dramatically changed the release probability for NM (Figure SI_4). Although the different NFs differed significantly by their powder dustiness, the Tier 1_NEP properties which predict their release potential from NEPs (Figure 1) were the same for each NEP matrix and independent of the embedded NF, as confirmed experimentally by three independent laboratories (Figure SI_4). As an exception CuO added to wood as a biocidal preservative differed from all other NEP case studies because CuO was attributed as "attached" only. We earlier showed that Cu ions have to be released for biocidal performance.[78]. Fragments released from NEPs consisting of the same *solid* matrix but different

embedded NFs have been shown to be similarly toxic for animals or humans,[37-40, 79] We verified this grouping hypothesis on the case study of automotive coatings containing CuPhthalocyanine.[42] Fragments from NEPs consisting of the same *solid* matrix but different *embedded* NFs (characterised in Tier 1_NEP) are also similar in their ecotoxicity.[26] Thus it appears that if the state of dispersion and content of a particular NF in the NEP, and also its intended use is the same, then the lifecycle-induced fragmentation is- similar in form, rate, toxicity and ecotoxicity. In our case studies, all Tier 1_NEP properties that could predict the rate and form of the release of fragments were the same for different nanoforms in the same solid matrix (see Figure SI_4), supporting a grouping according to lifecycle issues primarily by the matrix material and the intended use of the NEP. Other accompanying studies confirmed this grouping hypothesis specifically for the automotive coatings with different embedded nano- and non-nano-pigments, and for the lightweight materials with different embedded CNT, graphene, CB or Kaolin fillers. [80-82] Such a grouping is considered to be relevant by US-EPA and industrial downstream users but is not foreseen to be used in the context of European regulation e.g. REACH. ISO TC229 PG29 is currently working on the standardisation of this grouping in form of a decision tree. We found that the grouping by matrix and intended use only fails where the lifecycle process elicits catalytic activity, such as metal-based NFs during NEP incineration,[83] or where the NEP formation itself is a reactive process such as cement hydration, where the "matrix" crystallinity and porosity is controlled by the NFs, and whose formation consumes the NFs, which is the very purpose of the slag (GGBS) to replace CO_2-intensive clinker[84] or of nano SiO_2 in a "pozzolanic reaction".

Several of the case studies have been evaluated by other grouping frameworks before and focussed especially on OECD NM and on some of the pigments. Comparing e.g. the grouping by occupational (inhalation) safety of ZnO NM111, both the nanoGRAVUR framework and the DF4nanogrouping framework categorize NM111 in Tier 2 as "biosoluble" in the relevant medium and read across to the non-nano form. In contrast, Pigment Blue 15:1 (Cu-Phthalocyanine) was a false positive in the binary decision logic of the DF4nanogrouping framework, being "active" in Tier 2, but "*in vivo* passive" in a Tier 3 calibration. In the risk matrix of nanoGRAVUR[85] several criteria are combined and result in a "green" risk group, consistent with the "*in vivo* passive" Tier 3 calibration. Of note, the risk matrices that build on the current framework are only intended for non-regulatory use. In the current framework, fibres are still recognised as a specific group already in Tier 1, triggering adapted tests such as descriptors of the form that is released during the intended use. In future, benchmark materials (e.g. Mitsui NT7) and methods for rigidity are foreseen to enhance the robustness of the HARN category. In this context it has to be mentioned that the trigger values for fibres regarding human toxicity and ecotoxicity presumably differ and still need to be defined for ecotoxicity [25].

We did not test polymer particles among our case studies, but the tested pigments (DPP NFs and CuPhthalocyanine NFs) represent hydrophobic organic particles with low human toxicity[68] and low ecotoxicity.[25] No unique ecotoxicity[86, 87] or human toxicity[88] of polymer NFs was discovered, and hence properties beyond those of our framework should not be required to assess the similarity between polymer particles. There may be limits if the polymer NF is significantly soluble or swelling, such as hydrogels, for which not enough data exists to draw conclusion.

As noted by Gao & Lowry, extrinsic properties result from a more or less complex interplay of the surrounding medium and one or more intrinsic properties,[89] but some extrinsic properties are less sensitive than others towards variations of size, shape etc. of different NFs of the same substance. The assessment of dissolution and transformation by either the environmental or the human perspective typically assigns different NFs of the same substance to the same band, with rare exceptions. This property is determined more by the substance than by size or shape or crystallinity or organic surface treatment (we did not test true core-shell systems, which would certainly modulate the dissolution behavior).

We observed that, when different NFs of the same substance are evaluated by harmonized methods, they are *often* assigned to different bands of dustiness, *sometimes* to different bands in the dispersion stability and mobility in soils, *sometimes* to different bands in abiotic reactivity, but *very often* to different bands of the by descriptor of *in vitro* (NR8383) reactivity, which is constructed from several read-outs. The sensitivity of dustiness to different NFs is supported by literature, which furthermore supports the order-of-magnitude (decadic) bands.[58, 59, 90]

Not only for reactivity, our construction of descriptors and bands may in fact exaggerate dissimilarity: e.g. for CeO_2 reactivity, TiO_2 reactivity, GGBS reactivity or SiO_2 dispersion stability, the numerical values are similar (Table 3), but fall on different bands of the universally fixed order-of-magnitude (decadic) cutoffs. Instead of the risk-screening purpose, for which fixed bands are appropriate, the grouping purpose may be better served by floating band center values with a fixed order-of-magnitude (decadic) span. The floating bands would focus on similarity within a given candidate group of NFs but would still keep the evaluation transparent with independent assessment of each property. We may also have to reconsider the normalization of the NR8383 NOAEC by specific surface areas,[63] and the indexing by it, because –apart from partico-kinetics, cell contact and particle uptake into cells– there is no simple and apparent reason why the non-nano-form of $BaSO_4$ should be more reactive than the NF, or why the non-nano-form of DPP should be significantly more reactive than the NF. Possibly the NR8383 evaluation can be adapted to the same concept of order-of-magnitude (decadic) bands as used for most other descriptors. However, this demands careful and quantitative studies on particle uptake, especially for well dispersed (i.e. diffusing, not sedimenting) particles.

Any property descriptors that are constructed from several redundant assays need data reduction strategies. On the

example of „biological reactivity“, we combined abiotic reactivity (two assays, four read-outs) and one *in vitro* assay (another four read-outs), but our addition of „points “(Table 2) is not optimal and triggers numerous cases of „false positive dissimilarity“. False positives in Tier 2 require the user who wishes to use grouping to escalate to Tier 3 testing, only to find there that the NFs are similar, such that animal testing should have been prevented by a more efficient Tier 2.

In summary of the case study calibration by the *in vivo* studies and ecotox OECD studies, „ideal” similarity (i.e. different NFs are assigned to the same bands for *all* Tier 2 properties that are relevant for the grouping purpose by Table 1) resulted in false positives (apparent dissimilarity) in too many cases and would have triggered extensive animal testing. The more pragmatic weight-of-evidence approach (i.e., joint assessment of properties that are relevant for the grouping purpose by Table 1) is more appropriate to implement the stepwise ECHA process of hypothesis formulation, data gathering and hypothesis substantiation for regulatory grouping.[16].[16]

In this perspective the outstanding importance of benchmark materials becomes evident: We cannot assess the significance of any dissimilarity, if we do not know the dynamic range of that particular property. Benchmark materials span the dynamic range. We observe that for many properties the dynamic range spans about three to four orders of magnitude between the positive and negative controls (i.e. benchmark materials). This applies to size, aspect ratio, solubility, reactivity, dissolution rate, dustiness, ENM content, resilience of matrix, but not to homo-agglomeration or *invitro*-reactivity. Thus, order-of-magnitude (decadic) band ranges are appropriate for NFs based on experimental evidence, and this is fully in line with the GHS, where such a factor of 10 is often applied in the assessment of non-nano chemical substances. As discussed above, floating band center values would be the next step from risk-screening perspective towards a grouping by similarity (instead of categorization and banding). But, alternatively to bands, one may develop algorithms that compare the pairwise distance of materials in a multidimensional space spanned by the relevant properties (Table 1, Figure 1), and compare it to the distance of benchmark materials (Table 2, Table 5). Such concepts are explored in follow-up projects such as GRACIOUS (H2020) and InnoMat.Life (BMBF).

Conclusions

The methods developed or selected by nanoGRAVUR fill several gaps highlighted in the Steinhäuser & Sayre (2017) reviews and are useful to implement both the ECHA concept of grouping of *nanoforms* or sets thereof, as well as the EPA concept of *discrete forms*. Previous frameworks had a narrower focus on occupational or ecological hazard. The nanoGRAVUR framework serves three purposes of grouping for occupational, environmental and consumer safety. Depending on the purpose, different properties become relevant to assess the similarity. Grouping decisions can be made in the Tier 2 mostly based on extrinsic properties with quantitative bands that are order-of-magnitude (decadic) for many properties. Benchmark materials span the dynamic range, which in general crosses about three to four orders of magnitude.

Case studies include families of Fe_2O_3, SiO_2, Aluminosilicates, $BaSO_4$, CeO_2, organic pigment, ZnO, TiO_2 (nano)forms. We find that for some substances the biological similarities were high when (nano)forms only differed in morphology and particle size, specifically for NFs and non-nano-forms of SiO_2, $BaSO_4$, Kaolin, CeO_2, ZnO, organic pigments. In contrast, different Fe_2O_3 or TiO_2 (nano)forms differ more significantly. Surface treatments in the sense of the ECHA guidance[14] were tested on ZnO, TiO_2, SiO_2, and were found to modulate to a significant extent the dispersion stability and the reactivity, but only had transient influence on dissolution rates in the presented cases.[54]

We further observed that, when different NFs of the same substance are evaluated by harmonised methods, they are often assigned to different bands of dustiness, sometimes to different bands in the dispersion stability and mobility in soils, sometimes to different bands in abiotic reactivity, and very often to different bands by the descriptors of *in vitro* and abiotic reactivity. In contrast, the dissolution and transformation behaviour, measured by several approaches, was primarily determined by the substance.

The NFs were also integrated in nano-enabled products (NEPs) for automotive coatings, clinker-reduced cements, cosmetic sunscreen, lightweight polymers. Once incorporated into a certain NEP with solid matrix, all relevant properties that could predict and assess the rate and form of release were within an order-of-magnitude (decadic) band for different NFs in the same solid matrix; together with the state-of-the-art on (eco)toxicity of NEP fragments, our findings support a grouping of lifecycle issues primarily by intended use and NEP matrix.

The evaluation of the developed framework with a diverse set of case studies clearly showed its usability for grouping for different purposes – but also the limits: We still need to develop rules for read-across, to explore measures that quantify similarity across multiple descriptors. The nanoGRAVUR case studies, thoroughly tested by ISO standards and OECD guidelines, will be essential to validate any future frameworks.

Conflicts of interest

WW, JK, RL, BF are employees of companies producing and marketing nanomaterials.

Acknowledgements

This work was partially funded by the BMBF (German Federal Ministry of Education and Research) project nanoGRAVUR - Nanostructured materials – Grouping in view of worker, consumer and environmental safety and risk minimization (FKZ 03XP0002X). We gratefully acknowledge the work and contribution of the entire consortium to materials, data and concepts namely by Axel Horster (General Reinsurance); Nicole Neubauer, Emmanuel Ruggiero, Patrick Fiedler (BASF SE); Bianca Oeffling, Carsten Möhlmann (Institut für Arbeitsschutz

der DGUV); Michael Kirchner (Institut für Gefahrstoff-Forschung der BGRCI); Jürgen Schnekenburger (Westfälische Wilhelms-Universität Münster); Burkhard Stahlmecke (IUTA); Frederic Krabigell, Bernd Staubach (HDI Global SE); Tobias Teckentrup (Universität Duisburg-Essen); and by Umweltbundesamt (Dessau).

References

1. E. P. A. (EPA), *Federal Register*, 2017, **82**, 22088.
2. T. C. Iso, *ISO/AWI TS 80004-4*, 2010.
3. W. Wohlleben, C. Punckt, J. Aghassi-Hagmann, F. Siebers, F. Menzel, D. Esken, C.-P. Drexel, H. Zoz, H. U. Benz, A. Weier, M. Hitzler, A. I. Schäfer, L. D. Cola and E. A. Prasetyanto, in *Metrology and Standardization for Nanotechnology: Protocols and Industrial Innovations*, eds. E. Mansfield, D. L. Kaiser, D. Fujita and M. Van de Voorde, John Wiley & Sons, 2017, pp. 411-464.
4. W. Stark, P. Stoessel, W. Wohlleben and A. Hafner, *Chemical Society Reviews*, 2015, **44**, 5793-5805.
5. EC, Comission recommendation of 18 October 2011 on the definition of nanomaterial*, Official Journal of the European Union*, 2011.
6. K. Aschberger, H. Rauscher, H. Crutzen, K. Rasmussen, F. M. Christensen, B. Sokull-Klüttgen and H. Stamm, *JRC Science and Policy Reports*, 2017, DOI: 10.2788/3044
7. A. G. Oomen, K. G. Steinhäuser, E. A. J. Bleeker, F. van Broekhuizen, A. Sips, S. Dekkers, S. W. P. Wijnhoven and P. G. Sayre, *NanoImpact*, 2018, **9**, 1-13.
8. J. H. Arts, M. Hadi, A. M. Keene, R. Kreiling, D. Lyon, M. Maier, K. Michel, T. Petry, U. G. Sauer, D. Warheit, K. Wiench and R. Landsiedel, *Regul Toxicol Pharmacol*, 2014, **70**.
9. A. G. Oomen, E. A. Bleeker, P. M. Bos, F. van Broekhuizen, S. Gottardo, M. Groenewold, D. Hristozov, K. Hund-Rinke, M.-A. Irfan, A. Marcomini, W. J. Peijnenburg, K. RASMUSSEN, A. Sanchez Jimenez, J. Scott-Fordsmand, M. Van Tongeren, K. Wiench, W. Wohlleben and R. Landsiedel, *International journal of environmental research and public health*, 2015, **12**, 13415-13434.
10. N. Burden, K. Aschberger, Q. Chaudhry, M. J. Clift, S. H. Doak, P. Fowler, H. Johnston, R. Landsiedel, J. Rowland and V. Stone, *Nano Today*, 2017, **12**, 10-13.
11. H. Godwin, C. Nameth, D. Avery, L. L. Bergeson, D. Bernard, E. Beryt, W. Boyes, S. Brown, A. J. Clippinger, Y. Cohen, M. Doa, C. O. Hendren, P. Holden, K. Houck, A. B. Kane, F. Klaessig, T. Kodas, R. Landsiedel, I. Lynch, T. Malloy, M. B. Miller, J. Muller, G. Oberdorster, E. J. Petersen, R. C. Pleus, P. Sayre, V. Stone, K. M. Sullivan, J. Tentschert, P. Wallis and A. E. Nel, *ACS Nano*, 2015, **9**, 3409-3417.
12. A. E. Nel, E. Nasser, H. Godwin, D. Avery, T. Bahadori, L. Bergeson, E. Beryt, J. C. Bonner, D. Boverhof, J. Carter, V. Castranova, J. R. Deshazo, S. M. Hussain, A. B. Kane, F. Klaessig, E. Kuempel, M. Lafranconi, R. Landsiedel, T. Malloy, M. B. Miller, J. Morris, K. Moss, G. Oberdörster, K. Pinkerton, R. C. Pleus, J. A. Shatkin, R. Thomas, T. Tolaymat, A. Wang and J. Wong, *ACS Nano*, 2013, **7**.
13. ECHA, *How to prepare registration dossiers that cover nanoforms: best practices*, 2017, DOI: 10.2823/128306.
14. ECHA, *Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals*, 2017, DOI: 10.2823/884050.
15. E. D. Kuempel, V. Castranova, C. L. Geraci and P. A. Schulte, *J Nanopart Res*, 2012, **14**.
16. N. M. Drew, E. D. Kuempel, Y. Pei and F. Yang, *Regulatory Toxicology and Pharmacology*, 2017, **89**, 253-267.
17. J. H. E. Arts, M. Hadi, M.-A. Irfan, A. M. Keene, R. Kreiling, D. Lyon, M. Maier, K. Michel, T. Petry, U. G. Sauer, D. Warheit, K. Wiench, W. Wohlleben and R. Landsiedel, *Regulatory Toxicology and Pharmacology*, 2015, DOI: http://dx.doi.org/10.1016/j.yrtph.2015.03.007.
18. J. H. E. Arts, M.-A. Irfan, A. M. Keene, R. Kreiling, D. Lyon, M. Maier, K. Michel, N. Neubauer, T. Petry, U. G. Sauer, D. Warheit, K. Wiench, W. Wohlleben and R. Landsiedel, *Regulatory Toxicology and Pharmacology*, 2016, **76**, 234-261.
19. A. Gajewicz, *Environmental Science: Nano*, 2018, **5**, 408-421.
20. R. Landsiedel, L. Ma-Hock, K. Wiench, W. Wohlleben and U. G. Sauer, *Journal of Nanoparticle Research*, 2017, **19**, 171.
21. Z. Wang, M. G. Vijver and W. J. Peijnenburg, *Journal*, 2018.
22. K. Hund-Rinke, K. Schlich, D. Kühnel, B. Hellack, H. Kaminski and C. Nickel, *NanoImpact*, 2018, **9**, 52-60.
23. D. Kühnel, C. Nickel, E. van der Zalm, C. Kussatz, M. Herrchen, B. Meisterjahn and K. Hund-Rinke, *NanoImpact* **15**: 100173.
24. M. J. B. Amorim, S. Lin, K. Schlich, J. M. Navas, A. Brunelli, N. Neubauer, K. Vilsmeier, A. L. Costa, A. Gondikas, T. Xia, L. Galbis, E. Badetti, A. Marcomini, D. Hristozov, F. v. d. Kammer, K. Hund-Rinke, J. J. Scott-Fordsmand, A. Nel and W. Wohlleben, *Environmental Science & Technology*, 2018, DOI: 10.1021/acs.est.7b04122.
25. T. van Harmelen, E. K. Zondervan-van den Beuken, D. H. Brouwer, E. Kuijpers, W. Fransman, H. B. Buist, T. N. Ligthart, I. Hincapié, R. Hischier, I. Linkov, B. Nowack, J. Studer, L. Hilty and C. Som, *Environment International*, 2016, **91**, 150-160.
26. M. L. Fernandez-Cruz, D. Hernandez-Moreno, J. Catalan, R. K. Cross, H. Stockmann-Juvala, J. Cabellos, V. R. Lopes, M. Matzke, N. Ferraz, J. J. Izquierdo, J. M. Navas, M. Park, C. Svendsen and G. Janer, *Environmental Science: Nano*, 2018, **5**, 381-397.
27. L. Pizzol, D. Hristozov, A. Zabeo, G. Basei, W. Wohlleben, A. J. Koivisto, K. A. Jensen, W. Fransman, V. Stone and A. Marcomini, *Nanoimpact*, 2019, **13**, 26-36.
28. D. Hristozov, S. Gottardo, E. Semenzin, A. Oomen, P. Bos, W. Peijnenburg, M. van Tongeren, B. Nowack, N. Hunt and A. Brunelli, *Environment international*, 2016, **95**, 36-53.
29. ECHA, *Guidance on information requirements and chemical safety assessment*, 2012.
30. K. G. Steinhäuser and P. G. Sayre, *NanoImpact*, 2017, **7**, 66-74.
31. C. O. Hendren, G. V. Lowry, J. M. Unrine and M. R. Wiesner, *Science of The Total Environment*, 2015, **536**, 1029-1037.
32. H. Zhang, Z. Ji, T. Xia, H. Meng, C. Low-Kam, R. Liu, S. Pokhrel, S. Lin, X. Wang, Y. P. Liao, M. Wang, L. Li, R. Rallo,

R. Damoiseaux, D. Telesca, L. M+¤dler, Y. Cohen, J. I. Zink and A. E. Nel, *Acs Nano*, 2012, **6**, 4349-4368.
33. F. Abdolahpur Monikh, A. Praetorius, A. Schmid, P. Kozin, B. Meisterjahn, E. Makarova, T. Hofmann and F. von der Kammer, *NanoImpact*, 2018, **11**, 42-50.
34. R. C.-o. C. N. I. (RCC), *Work Area 3: Risk Assessment-Risk Management*, 2014.
35. W. Wohlleben, S. Brill, M. Meier, M. Mertler, G. Cox, S. Hirth, B. von Vacano, V. Strauss, S. Treumann, K. Wiench, L. Ma-Hock and R. Landsiedel, *Small*, 2011, **7**, 2384 - 2395.
36. W. Wohlleben, M. W. Meier, S. Vogel, R. Landsiedel, G. Cox, S. Hirth and Ž. Tomović, *Nanoscale*, 2013, **5**, 369-380.
37. A. T. Saber, A. Mortensen, J. Szarek, I. K. Koponen, M. Levin, N. R. Jacobsen, M. E. Pozzebon, S. P. Mucelli, D. G. Rickerby, K. Kling, R. Atluri, A. M. Madsen, P. Jackson, Z. O. Kyjovska, U. Vogel, K. A. Jensen and H. Wallin, *Particle and Fibre Toxicology*, 2016, **13**, 1-20.
38. A. Saber, N. Jacobsen, A. Mortensen, J. Szarek, P. Jackson, A. Madsen, K. Jensen, I. Koponen, G. Brunborg, K. Gutzkow, U. Vogel and H. Wallin, *Part.Fibre Toxicol.*, 2012, **9**, 4.
39. A. Saber, I. Koponen, K. Jensen, N. Jacobsen, L. Mikkelsen, P. Moller, S. Loft, U. Vogel and H. Wallin, *Nanotoxicology*, 2012, **6**, 776 - 788.
40. C. Pang, N. Neubauer, M. Boyles, D. Brown, N. Kanase, D. Hristozov, T. Fernandes, V. Stone, W. Wohlleben and A. Marcomini, *NanoImpact*, 2017, **7**, 75-83.
41. J. J. Scott-Fordsmand, J. M. Navas, K. Hund-Rinke, B. Nowack and M. J. B. Amorim, *Nano Today*, 2017, DOI: https://doi.org/10.1016/j.nantod.2017.09.002.
42. E. Kuempel, V. Castranova, C. Geraci and P. Schulte, *J. Nanopart Res*, 2012, **14**, 1029.
43. AGS, *Committee on Hazardous Substances*, 2013, **527**.
44. K. Bhattacharya, G. Kiliç, P. M. Costa and B. Fadeel, *Nanotoxicology*, 2017, **11**, 809-826.
45. V. Castagnola, J. Cookman, J. M. de Araujo, E. Polo, Q. Cai, C. P. Silveira, Z. Krpetic, Y. Yan, L. Boselli and K. A. Dawson, *Nanoscale Horizons*, 2017, **2**, 187-198.
46. Q. Mu, G. Jiang, L. Chen, H. Zhou, D. Fourches, A. Tropsha and B. Yan, *Chemical reviews*, 2014, **114**, 7740-7781.
47. F. Babick, J. Mielke, W. Wohlleben, S. Weigel and V.-D. Hodoroaba, *Journal of Nanoparticle Research*, 2016, **18**, 1-40.
48. W. Wohlleben, J. Mielke, A. Bianchin, A. Ghanem, H. Freiberger, H. Rauscher, M. Gemeinert and V.-D. Hodoroaba, *Journal of Nanoparticle Research*, 2017, **19**, 61.
49. D. Mehn, I. M. Rio-Echevarria, D. Gilliland, M. Kaiser, K. Vilsmeier, P. Schuck and W. Wohlleben, *NanoImpact*, 2018, **10**, 87-96.
50. OECD, *ISO/TR 19057:2017*, 2017.
51. J. Koltermann-Jülly, J. G. Keller, A. Vennemann, K. Werle, P. Müller, L. Ma-Hock, R. Landsiedel, M. Wiemann and W. Wohlleben, *NanoImpact*, 2018, DOI: https://doi.org/10.1016/j.impact.2018.08.005.
52. G. Oberdörster and T. A. J. Kuhlbusch, *NanoImpact*, 2018, **10**, 38-60.
53. W. Wohlleben, C. Kingston, J. Carter, E. Sahle-Demessie, S. Vázquez-Campos, B. Acrey, C.-Y. Chen, E. Walton, H. Egenolf, P. Müller and R. Zepp, *Carbon*, 2017, **113**, 346-360.
54. T. Schneider and K. A. Jensen, *Annals of Occupational Hygiene*, 2007, **52**, 23-34.
55. D. Broßell, E. Heunisch, A. Meyer-Plath, D. Bäger, V. Bachmann, K. Kämpf, N. Dziurowitz, C. Thim, D. Wenzlaff and J. Schumann, *Powder technology*, 2019, **342**, 491-508.
56. C. Dazon, O. Witschger, S. Bau, R. Payet, K. Beugnon, G. Petit, T. Garin and L. Martinon, 2017.
57. I. Pensis, J. Mareels, D. Dahmann and D. Mark, *Annals of occupational hygiene*, 2009, **54**, 204-216.
58. D. Göhler and M. Stintz, *Journal of Physics: Conference Series*, 2015, **617**, 012029.
59. B. Hellack, C. Nickel, C. Albrecht, T. A. J. Kuhlbusch, S. Boland, A. Baeza-Squiban, W. Wohlleben and R. P. F. Schins, *Environmental Science: Nano*, 2017, **4**, 1920-1934.
60. M. Wiemann, A. Vennemann, U. G. Sauer, K. Wiench, L. Ma-Hock and R. Landsiedel, *Journal of Nanobiotechnology*, 2016, **14**, 1-27.
61. A. B. Kane, R. H. Hurt and H. Gao, *Toxicology and applied pharmacology*, 2018, **361**, 68-80.
62. T. Hofmann, L. Ma-Hock, V. Strauss, S. Treumann, M. Rey Moreno, N. Neubauer, W. Wohlleben, S. Gröters, K. Wiench, U. Veith, W. Teubner, B. van Ravenzwaay and R. Landsiedel, *Inhalation Toxicology*, 2016, **28**, 463-479.
63. R. Landsiedel, L. Ma-Hock, T. Hofmann, M. Wiemann, V. Strauss, S. Treumann, W. Wohlleben, S. Groeters, K. Wiench and B. Ravenzwaay, *Part Fibre Toxicol*, 2014, **11**.
64. L. Ma-Hock, V. Strauss, S. Treumann, K. Kuttler, W. Wohlleben, T. Hofmann, S. Groters, K. Wiench, R. B. van and R. Landsiedel, *Part Fibre.Toxicol.*, 2013, **10**, 23.
65. J. Keller, W. Wohlleben, L. Ma-Hock, V. Strauss, S. Gröters, K. Küttler, K. Wiench, C. Herden, G. Oberdörster and B. van Ravenzwaay, *Archives of toxicology*, 2014, **88**, 2033-2059.
66. R. Landsiedel, L. Ma-Hock, A. Kroll, D. Hahn, J. Schnekenburger, K. Wiench and W. Wohlleben, *Adv.Mater.*, 2010, **22**, 2601-2627.
67. I. Gosens, F. R. Cassee, M. Zanella, L. Manodori, A. Brunelli, A. L. Costa, B. G. H. Bokkers, W. H. de Jong, D. Brown, D. Hristozov and V. Stone, *Nanotoxicology*, 2016, **10**, 1084-1095.
68. J. H. E. Arts, H. Muijser, E. Duistermaat, K. Junker and C. F. Kuper, *Food and Chemical Toxicology*, 2007, **45**, 1856-1867.
69. R. Landsiedel, J. Schnekenburger, F. Alessandrini, A. Haase, A. Luch, L. Ma-Hock and M. Wiemann, *Journal*, 2014, DOI: 10.2314/GBV:82774322X.
70. W. Wohlleben, M. D. Driessen, S. Raesch, U. F. Schaefer, C. Schulze, B. v. Vacano, A. Vennemann, M. Wiemann, C. A. Ruge, H. Platsch, S. Mues, R. Ossig, J. M. Tomm, J. Schnekenburger, T. A. J. Kuhlbusch, A. Luch, C.-M. Lehr and A. Haase, *Nanotoxicology*, 2016, DOI: 10.3109/17435390.2016.1155671, 1-11.
71. L. Chen, C. Wang, S. Yang, X. Guan, Q. Zhang, M. Shi, S.-T. Yang, C. Chen and X.-L. Chang, *Environmental Science: Nano*, 2019, **6**, 1077-1088.
72. D. Pantano, N. Neubauer, J. Navratilova, L. Scifo, C. Civardi, V. Stone, F. von der Kammer, P. Müller, M. S. Sobrido, B. Angeletti, J. Rose and W. Wohlleben, *Environmental Science & Technology*, 2018, DOI: 10.1021/acs.est.7b04130.
73. J. P. Kaiser, M. Roesslein, L. Diener and P. Wick, *PLOS ONE*, 2013, **8**, e83215.

74. E. Ruggiero, K. Vilsmeier, P. Mueller, S. Pulbere and W. Wohlleben, Manuscript submitted 2019.
75. R. Lankone, E. Ruggiero, D. Goodwin, L. P. Sung and W. Wendel, Manuscript submitted 2019.
76. Z. Richard, S.-D. Endalkachew, H. Hsin-Se, A. Brad, H. Changseok, D. M. J. B., V. Klaus, E. Ruggiero and W. Wendel, Manuscript in preparation 2019.
77. D. Singh, W. Wohlleben, R. D. L. T. Roche, J. C. White and P. Demokritou, *NanoImpact*, 2019, **13**, 44-55.
78. B. Funk, D. Göhler, B. Sachsenhauser, M. Stintz, B. Stahlmecke, B. A. Johnson and W. Wohlleben, *Environmental Science: Nano*, 2019, **6**, 1443-1456.
79. T. A. J. Kuhlbusch, S. W. P. Wijnhoven and A. Haase, *NanoImpact*, 2018, **10**, 11-25.
80. V. Adam, T. Yang and B. Nowack, *Environmental Toxicology and Chemistry*, 2019, **38**, 436-447.
81. M. Hauser, G. Li and B. Nowack, *Journal of nanobiotechnology*, 2019, **17**, 56.
82. R. Lehner, C. Weder, A. Petri-Fink and B. Rothen-Rutishauser, *Environmental science & technology*, 2019, **53**, 1748-1765.
83. X. Gao and G. V. Lowry, *NanoImpact*, 2018, **9**, 14-30.
84. T. Schneider and K. A. Jensen, *Annals of Occupational Hygiene*, 2008, **52**, 23-34.

