Organisation for Economic Co-operation and Development

ENV/JM/MONO(2018)28

Unclassified

English - Or. English

21 September 2018

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

Cancels & replaces the same document of 19 September 2018

CASE STUDY ON GROUPING AND READ-ACROSS FOR NANOMATERIALS – GENOTOXICITY OF NANO-TIO2

Series on Testing and Assessment
No. 292

JT03435966

OECD Environment, Health and Safety Publications

Series on Testing and Assessment

No. 292

CASE STUDY ON GROUPING AND READ-ACROSS FOR NANOMATERIALS GENOTOXICITY OF NANO- TiO2

IOMC

INTER-ORGANIZATION PROGRAMME FOR THE SOUND MANAGEMENT OF CHEMICALS

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

Environment Directorate

ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT

Paris 2018

About the OECD

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

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

This publication is available electronically, at no charge.

Also published in the Series on testing and Assessment link

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

or contact:

OECD Environment Directorate,
Environment, Health and Safety Division
2, rue André-Pascal
75775 Paris cedex 16
France
Fax : (33-1) 44 30 61 80
E-mail : ehscont@oecd.org

FOREWORD

OECD member countries have been making efforts to expand the use of alternative methods in assessing chemicals. The OECD has been developing guidance documents and tools for the use of alternative methods such as (Q)SAR, chemical categories and Adverse Outcome Pathways (AOPs) as a part of Integrated Approaches for Testing and Assessment (IATA). There is a need for the investigation of the practical applicability of these methods/tools for different aspects of regulatory decision-making, and to build upon case studies and assessment experience across jurisdictions.

The objective of the IATA Case Studies Project is to increase experience with the use of IATA by developing case studies, which constitute examples of predictions that are fit for regulatory use. The aim is to create common understanding of using novel methodologies and the generation of considerations/guidance stemming from these case studies.

This case study was developed by European Union / Joint Research Centre (JRC) for illustrating practical use of IATA and submitted to the 2017 review cycle of the IATA Case Studies Project. This case study was reviewed by the project team. The document was endorsed at the 2nd meeting of the Working Party on Hazard Assessment in June 2018.

The following three case studies were also reviewed in the project in 2017 and are published with this case study:

1. CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA) FOR ESTROGENICITY OF THE SUBSTITUTED PHENOLS, ENV/JM/MONO(2018)26, Series on Testing & Assessment No. 290.
2. PRIORITISATION OF CHEMICALS USING THE INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA)-BASED ECOLOGICAL RISK CLASSIFICATION, ENV/JM/MONO(2018)27, Series on Testing & Assessment No. 291.
3. A CASE STUDY ON THE USE OF INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT FOR SUB-CHRONIC REPEATED-DOSE TOXICITY OF SIMPLE ARYL ALCOHOL ALKYL CARBOXYLIC ESTERS: READ-ACROSS, ENV/JM/MONO(2018)29, Series on Testing & Assessment No. 293.

These case studies are illustrative examples, and their publication as OECD monographs does not translate into direct acceptance of the methodologies for regulatory purposes across OECD countries. In addition, these cases studies should not be interpreted as official regulatory decisions made by the authoring member countries.

A considerations document summarizing the learnings and lessons of the review experience of the case studies is published with the case studies:

REPORT ON CONSIDERATIONS FROM CASE STUDIES ON INTEGRATED APPROACHES FOR TESTING AND ASSESSMENT (IATA) -Third Review Cycle (2017) - ENV/JM/MONO(2018)25, Series on Testing & Assessment No. 289.

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

DISCLAIMER

The report presents the findings and conclusions of the authors, but does not represent an official view of the JRC or other European Commission services.

SUMMARY

The case study is applying the workflow for grouping and read-across proposed in the ECHA REACH guidance update for nanomaterials (NMs) and exploring the extent to which ECHA's Read-Across Assessment Framework (RAAF) is applicable to nanoforms for identifying sources of uncertainty associated with the read-across. The purpose of the read-across exercise as such was to determine the genotoxic hazard potential of two nano-TiO_2 target substances based on *in vitro* comet assay results from other TiO_2 nanoforms.

The dataset of the analogues includes six NMs with different properties: different primary and crystallite sizes (5-100nm), different crystalline types (anatase and rutile) and surface characteristics (some are coated and some uncoated). The grouping hypothesis derived from the collected physicochemical characteristics and *in vitro* comet assay results is that nano-TiO_2 in its uncoated form may damage DNA, but this potential would be masked by the presence of non-reactive coating and large amounts of impurities on the surface of the NM. A set of chemoinformatics tools ─ hierarchical clustering, principal component analysis, variable selection with random forest ─ was used to identify the (physicochemical) properties that differentiate the analogues, determine their similarity and that may drive genotoxicity. The findings were used to verify and strengthen the grouping hypothesis and the link between NM properties and the effect considered. The outcome of the *in vitro* comet assay was predicted for two target TiO_2 NMs, negative for the coated one, and positive for the uncoated one. These results were verified by experimental literature data.

Overall the case study shows the practical application of the ECHA guidance for grouping and read-across of NMs, which was slightly adapted for the case study. Furthermore, the usefulness of cheminformatics techniques to support the grouping hypothesis by identifying the differences between nanoforms and by supporting the weight of evidence was demonstrated. The ECHA RAAF was successfully applied for evaluating the confidence in the read-across argumentation of NMs. Some nanospecific issues were identified for further specification of the RAAF for NMS, in particular the concept of similarity which cannot be based only on structural similarity for NMs.

Table of contents

INTRODUCTION

The aim of the read-across exercise in the case study is to determine the genotoxic hazard potential of two nano-TiO_2 target substances, by reading across *in vitro* comet assay results.

The case study is also evaluating the applicability of the workflow for grouping and read-across proposed in the REACH guidance update for nanomaterials (NMs) (ECHA 2017a) and identifying sources of uncertainty associated with the read-across, exploring the extent to which ECHA's Read-Across Assessment Framework (RAAF) (ECHA 2017b) captures the different sources of uncertainty for nanoforms.

The workflow for grouping and read-across proposed in the draft REACH guidance for nanomaterials as shown in Figure 1 will be referred to throughout the case study.

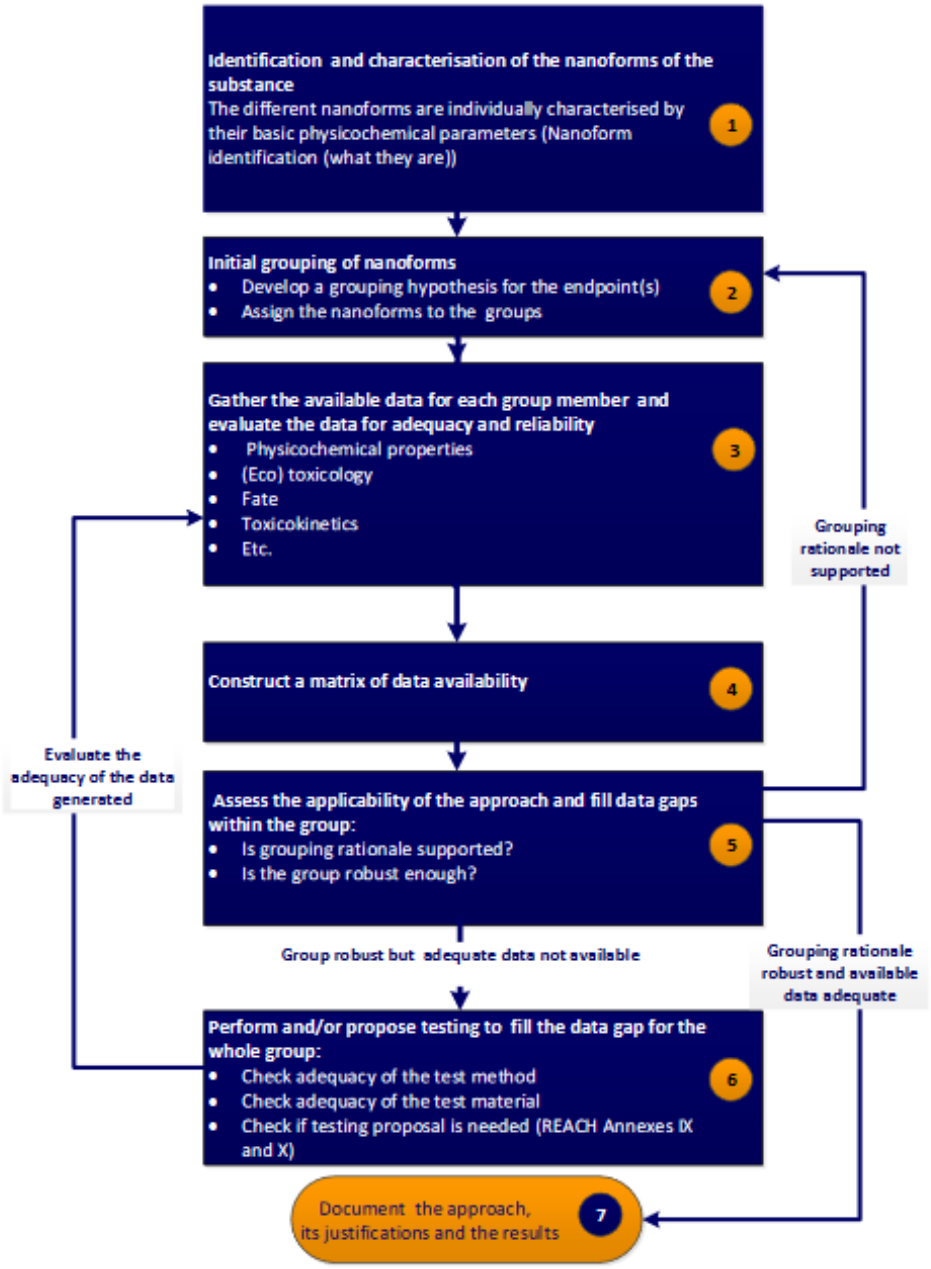

Figure 1. ECHA read-across workflow the case study was based on (ECHA 2017a).

1. PURPOSE

1.1. Purpose of the read-across

The aim of the read-across exercise is to determine the genotoxic hazard potential of two target nanoTiO_2 nanomaterials, based on *in vitro* comet assay results of other nano-TiO_2 analogues.

The work was carried out to explore the practical process of grouping and read-across between nanoforms, with a view to sharing the lessons learned about the overall process. Thus, the conclusions obtained for specific substances should not be regarded as recommendations for regulatory action.

1.2. Target chemicals

Identification and characterisation of the nanoforms of the target substances

In this case study, the result of the comet assay for TiO_2 Rutile (R) nano (Sigma 637262) and TiO_2 Anatase (A) nano (Sigma 637254) will be the data gaps to be filled by read-across.

The first step in the read-across workflow (see Figure 1) is the identification and appropriate characterisation of the nanoforms of the substance (ECHA 2017a). According to ECHA draft recommendations for the definition of nanoforms (ECHA 2017c), other than the composition (including impurities), the minimum requirements to register a nanomaterial for REACH are:

- Particle size (in one or more dimensions)
- Particle shape, e.g. spheroidal-like, high aspect ratio (≥5:1, nanotubes, nanorods), bidimensional (flakes or platelets), other (mixtures)
- Surface chemistry (chemical identity)

Table 1. Physicochemical properties of the nanomaterials used as target substances in the case study. Data obtained from (Guichard et al. 2012).

Properties	TiO_2 R nano	TiO_2 A nano
Crystal type	Rutile	Anatase
Total non-TiO_2 content (including coating and impurities) (% w/w)	13	0.50
Surface chemistry (as declared by manufacturer)	SiO_2 Na_2SO_4 SO_4	uncoated
Surface coating (%)	11	0
Primary particle diameter (nm)	10 nm diameter 62 nm length	14
Shape	Rod	Sphere
Specific surface area (m^2/g)	149	177

The information for the target NMs used in the present case study was extracted from Guichard et al. (2012) and presented in Table 1.

According to the physicochemical properties, the two target materials consist of a nanopowder with particles of either rutile or anatase TiO_2, specific surface area of 149 and 177 m^2/g, different levels of non-TiO_2 content (one is uncoated and has 99.5% purity, the coated one has 87% purity). These NM have been tested for cytotoxicity and genotoxicity in Syrian hamster embryo cells by Guichard et al. (2012). The reporting of results includes some physicochemical analysis that will be used to determine analogues, and the test results will be used to validate the read-across prediction.

The analysis of the physicochemical properties of the target substances shows that the measured ones are slightly different from those reported by the manufacturer. For instance, Guichard et al. (2012) found for TiO_2 R nano 11% of impurities corresponding mainly to SiO_2 (the manufacturer declared up to 5% of SiO_2 as surface coating), the measured particle size to corresponded to a rod of 62±10 x 24±2 nm (the manufacturer declared 40x10nm [1]), and the surface area to 177m^2/g (the manufacturer declared 50m^2/g). For the purpose of this study it is assumed that the substance used by Guichard et al. (2012) in their experiments really corresponds to the coated TiO_2 manufactured by Sigma. It is not clear though where is the limit to consider that two substances are the same.

1.3. Endpoint

The endpoint considered is genotoxic hazard as determined by the *in vitro* comet assay.

[1] According to details provided in the NM Aldrich catalogue http://www.sigmaaldrich.com/catalog/product/aldrich/637262?lang=it&region=IT

2. SOURCE CHEMICALS/CATEGORY MEMBERS

2.1. Identification and selection of source chemicals

To identify analogues of the target NM, the initial sources for collecting information on physicochemical properties and on toxicological endpoints were the Scientific Committee on Consumer Safety (SCCS) report and the OECD WPMN dossier on nano-TiO_2 (SCCS 2014, OECD 2015). The information obtained from these sources was used to construct a table with 53 nano-TiO_2. Unfortunately, most of these analogues could not be identified because the reported data (mainly in the SCCS report) had been anonymised, and the information reported did not include systematic physicochemical characterisation that could be used to identify them.

It was therefore decided to consider only the nanoforms that were identified properly by means of fundamental parameters such as solubility, hydrophobicity, zeta potential, and dispersability, which are among the properties that need to be taken into account for read-across of NMs according to ECHA 2017a. This led to a dataset with six nano-TiO_2 materials whose data were mainly obtained from the OECD dossier (from the version published online in March 2016).

2.2. Characterisation of the analogue nanoforms

This set of six TiO_2 nanoforms will be used as source substances (analogues) to predict via read-across the genotoxic potential of the two target nano-TiO_2, considering the *in vitro* comet assay.

The six nanoforms differ in their size (from 7 to 117 nm), coating (two of them are declared coated by the manufacturer and the others are declared without a coating), crystal type (anatase and rutile) and hydrophobicity (hydrophobic and hydrophilic). For an overview of physicochemical characterisation of these TiO_2, see Rasmussen et al. (2014). The physicochemical properties relevant for the identification ("what they are") of the six analogues are shown in Table 2.

Table 2. Physicochemical properties of the analogues obtained from the OECD dossiers (OECD 2015, downloaded March 2016).

Property	NM100	NM101	NM102	NM103	NM104	NM105
Crystal type	Anatase	Anatase	Anatase	Rutile	Rutile	83% anatase 17% rutile
Other info	Dry-milled	Semiconductor catalyst used in photocatalytic process	photocatalytic	hydrophobic	hydrophilic	-
Total non-TiO_2 content including coating and impurities (% w/w)	1.5	9	5	11	11	0.11
Surface chemistry (as declared by manufacturer)	uncoated	uncoated	uncoated	Al_2O_3 and SiO_2	Al_2O_3 and SiO_2	uncoated
Surface coating (% w/w)	0	0	0	8	8	0
Organic matter (%)	0	8	0	2	2	0
Primary particle diameter (TEM) (nm)	93 ± 23	5 ± 1	22 ± 10	24 ± 2	24 ± 2	20 ± 3
Crystallite size (XRD) (nm) [a]	117 ± 40	7 ± 2	24 ± 5	24 ± 4	25 ± 4	22 ± 5
Particle Size Distribution average (nm) [b]	210 ± 10	278	440 ± 37	135 ± 25	145 ± 35	177 ± 39
Shape	Spheroidal	Spheroidal	Spheroidal	Spheroidal	Spheroidal	Spheroidal
Aspect ratio	1.53	1.53	1.53	1.7	1.53	1.36
Specific surface area (m^2/g)	9	242 ± 73	77 ± 10	54 ± 4	54 ± 2	47 ± 0.5

a Values averaged from different instruments and principles (Peak fit, TOPAS, Fullprof, Scherrer eq., TOPAS, IB, TOPAS FWHM).
b Values averaged from inductively-coupled plasma mass spectrometry (ICP-MS) and dynamic light scattering (DLS) experiments.
c Values averaged from (ultra-) small-angle X-ray scattering (SAXS/USAXS) and BET (Brunauer, Emmett and Teller).

Impurities are important in identifying NMs; they are defined as "an unintended constituent present in a substance as manufactured" (ECHA 2017d); for NMs this is a different information compared to the surface coating that consists in the surface chemistry purposely added to the NM. The measurement of the elements present on the surface of the NM does not distinguish between the two, and hence the "Total non-TiO2 content including coating and impurities (% w/w)" is introduced in Table 2 as the measure for the total elements detected other than the core material. Thus, this measure includes both the impurities and the coating, which is separately declared by the supplier and is also reported separately in the dataset and in Table 2 as "Surface chemistry (as declared by manufacturer)", and "Surface coating (%)" indicating the quantity of coating with respect to total weight of the NM. NM-103 and NM-104 are declared coated and in fact the amount of total non-TiO2 content is >10%. NM-101 has quite a high content of impurities if compared to the uncoated NM-100, NM-102 and NM-105. The NM-101 "Total non-TiO_2 content including coating and impurities (% w/w)" is similar to that of NM-103 and NM-104.

The table shows one of the current problems of NMs. Some of the properties like crystallite size were obtained averaging different measures provided by different laboratories. In the particular case of NM-100, 117nm is the average of 141, 61, 168, and 100nm; some of them corresponding to triplicate measures with very low standard deviation. Similarly, the particle size diameter was obtained from two different measures, one of 70 ± 20nm and another one of 116.9 ± 36.9nm. NM-100 is a dry-milled

NM and the variability in the measurements of its particle diameter and crystallite size just indicate that it does not correspond to a mono-dispersed substance but that it contains a range of NM from 61-168nm. This will not represent a major problem in this case study as it would in an *in vivo* experiment, in which the size could significantly affect the exposure/distribution of the NM. Since the endpoint of interest in this case study is *in vitro* genotoxicity, the primary size of the NMs (i.e. not aggregated or agglomerated) will not affect the exposure. However, this can have big implications in *in vivo* endpoints in which the size of the particles may significantly affect the biodistribution (kinetics) of particles throughout the body and organs (e.g. lungs). The fact that the particle size diameter and crystallite size are very similar or that the latter is even larger is unexpected as it would indicate that the NMs are composed of just one crystal, which although possible is rather uncommon. Another factor that can help explaining this issue is the fact that the particle size diameter was measured with transmission electron microscopy (TEM) and the crystallite size by X-ray diffraction (XRD). In this case, the latter would be less accurate as the authors of the Nanogenotox report acknowledged having some problems fitting the data to some of the equations.

3. DATA GATHERING

Deviating slightly from the suggested order of the workflow in the ECHA guidance (ECHA 2017a), the next step was to gather data for the group members and evaluate the data for adequacy and reliability (step 3 in the ECHA workflow)(ECHA 2017). The collection of data and formulation of the initial grouping hypothesis (step 2 in the ECHA workflow) was experienced as an interactive process, since the hypothesis depended strongly on finding the relation between physicochemical properties and considered effects and thus on the available data. For a clearer description of the case study, the steps will be reported sequentially in the following.

3.1. Data gathering approach

The dataset was built containing two sets of information. Two clearly differentiated blocks of information were considered:

a) physicochemical characterisation, fundamental behaviour and reactivity of the identified NMs

b) toxicological data of relevant REACH endpoints/assays.

The choice of properties to capture in the database was informed by the templates proposed by Schultz et al. (Schultz et al. 2015), which were developed to help in assessing and reporting similarity assumptions in the context of chemistry, toxicokinetics and toxicodynamics, and thereby supporting read-across. Since the case studies presented here correspond to NMs, the templates were adapted by taking into account physicochemical parameters specific to NMs. These are identified in the REACH guidance update for NMs (ECHA 2017a) as key physicochemical parameters following a proposal by ITS-Nano (Stone et al. 2014). Other physicochemical parameters were taken into account. These parameters are either included in the OECD harmonised templates or considered relevant in the scientific literature.

The following physicochemical characteristics were collected in the dataset:

1. **What they are**: Name, JRC Nanomaterials Repository number, Chemical composition, Impurities, Crystal type, Crystal size, Surface coating, Porosity, Basic morphology, Primary particle diameter, Average particle diameter, Average length (TEM), Aspect ratio, Particle size distribution, Pour density (weighing), Specific surface area, Volume specific surface area

2. **Where they go:** Agglomeration, Hamaker constant, Dustiness, $\log K_{OW}$, Hydrophobicity, Biodurability, Dispersability, Stability of the dispersion, (Bio)persistence, Redox potential, Zeta potential, Isoelectric point, Abiotic transformation, Toxicokinetics, Metabolic products, Steric hindrance

3. **What they do**: Conduction band, Radical formation potential, Catalytic activity, Photocatalytic activity, Hydrolysis, Protein binding, Electrophilicity/Nucleophilicity (electrophilicity index), Physical hazards (flammability, autoflammability and explosiveness), Dissociation constant.

All the listed properties were searched for during data collection, but some were not used in developing the case studies because of lack of information or low data reliability.

To compile the Excel spreadsheets (the dataset) for the selected case studies, information was scrutinised from the following sources:

- Data available from public OECD dossiers (OECD 2015)
- REACH registration dossiers as of March 2016
- report on nano-TiO_2 proposal for classification from ANSES (ANSES 2016),
- JRC Repository (Rasmussen et al. 2014).
- Nanogenotox FP7 project (Nanogenotox 2012).

3.2. Data collected

Physicochemical parameters

Available information was collected considering the list of physicochemical properties reported above.

The total non-TiO_2 content of the analogues varies from 0.11% to 11%, where the highest values are justified by the presence of coating. NM-103 and NM-104 contain 6% of Al_2O_3 6 and 2% of organic functionalisation (silanes and dimethicone for NM-103, making it hydrophobic; tetramethyl silicate; glycerol; silanes; hexadecanoic acid, methyl ester, octadecaonic acid for NM-104 making it hydrophilic). NM-101 with 9% of non-TiO_2 content of organic composition (silane, Hexadecanoic acid, methyl ester, octadecaonic acid) could be considered to have coating although this nanoform was not declared coated by the manufacturer (Birkedal et al. 2012). This difference is reflected in Table 2, where the presence of surface coating is represented by its %w/w and where the Total non-TiO_2 content accounts for the amount of matter that is not TiO_2, thus including coating and impurities.

Since there was abundant information on the particle size distribution in biological media submitted in the OECD dossiers, this was also considered in the reported dataset.

The dependence of the NM properties on the environment (fate) is addressed in the dataset by including physicochemical properties measured in different media. For example, particle size distribution, zeta potential, polydispersity index are measured in Milli-Q water, in Dulbecco's modified eagle medium with and without L-glutamine, in foetal bovine serum and phosphate-buffered saline, whereas solubility and redox potential are measured in Gamble's solution (representing a lung fluid) and Caco2 medium (representing the intestinal environment). Dispersions were either untreated or underwent 1 minute probe sonication or 20 minutes ultrasound bath sonication. Inputs on biodurability were determined as the amount of Ti, Al, and Si dissolved in bovine serum albumin (BSA), Gamble's solution, and Caco2 solution after 24h exposure (Jensen et al. 2013).

The dataset of physicochemical parameters for the considered analogues, containing the data that was used in the analyses, is shown in Table 5 below. This data was analysed to identify a set of properties that can be used to define (structurally) similar NMs as well as to identify those physicochemical properties that can help in predicting by read-

across the results of an *in vitro* comet assay. The numerical values in Table 5 are deducted from a detailed data analysis and treatment (see Worth et al 2017, Appendices VIII and IX). A hierarchical clustering of the transposed matrixes of particle size distribution, zeta potential and polydispersity index (PdI) was used to select 4 properties for each group that were representative of the group. This step was necessary to avoid having as properties Zeta potential measured with Dulbecco's modified eagle medium (DMEM) and phosphate-buffered saline (PBS) 1% and Zeta potential measured with DMEM and PBS 5%. All the detail of how Table 5 was derived can be found in Worth et al 2017, Appendices VIII and IX. The full initial dataset is listed in Table A1 in the Annex.

Toxicological information

The analogues are quite data rich because they have been investigated in several FP7 projects and the WPMN testing program. The information available for each NM (NanoGenoTox Joint Action 2013; OECD 2015; SCCS 2014; Schröder et al. 2014; Sellers et al. 2015) was restructured starting from the OECD dossiers to build a complete dataset including both physicochemical information and toxicological information (genotoxicity) for NM-100, NM-101, NM-102, NM-103, NM-104, NM-105.

Results of the genotoxicity studies collected initially from these sources are included in Table 3. The table shows that there are *in vitro* comet assay results for all the identified analogues and they exhibit both positive and negative results. The results in the table report the positive tests over the total tests available. All the genotoxicity tests performed in the Nanogenotox WP6 were taken into consideration for the purpose of this case study. The testing protocol included a post-treatment of cythochalasin B that was added 6 h after the start of the treatment for all cell lines except CaCO2 where cytochalasin B was added after 24 h, so it can be assumed that this protocol applied the adaptation of the OECD TG 473 as recommended in Rasmussen et al. (2016). However, only the *in vitro* comet assay was performed for the six analogues. This is the considered endpoint for the case study. It should be noted that no OECD TG is available for this test.

ANSES (2016) performed a literature search on available genotoxicity case studies published in the period 2010-2015, reporting also registration dossiers collected by ECHA. Studies identified by ANSES that were missing in the original dataset were included and a further literature search was carried out to include other available studies. A reliability assessment of the studies was made according to the criteria identified by ANSES and reported here.

In vitro studies were considered reliable if:

1. The NMs were characterized (at least size, crystallinity and coating) and a description of the dispersed materials was provided (particle size distribution, zeta potential, polydispersity index)

2. The NM uptake was observed and/or cytotoxicity was tested

3. Positive and negative controls were considered, and replicates were included.

The *in vivo* studies were considered reliable if conditions 1 and 3 above applied; negative results were taken into account only when it was proven that the nanoparticles

had reached the organ investigated (condition 2). In *in vitro* studies, this could be confirmed with data on uptake or if cytotoxicity was detected.

In the data collection the micronucleus, comet and chromosomal aberration tests are reported because the current OECD test guidelines for these tests are considered applicable to NMs. Although some efforts are being done to evaluate the *in vitro* comet assay (Golbamaki et al. 2015; Azqueta & Dusinska 2015), an OECD test guideline is not available. This endpoint was considered however in the data collection because of larger availability of studies. Previous reviews on genotoxicity tests applied to NMs claimed that the comet and micronucleus assays are the most commonly used tests in the field (Golbamaki et al. 2015), and the search results confirm this.

Table 3 shows the results of the tests reported in the collected literature; the studies taken into consideration include the test results collected initially as reported above. The results are shown as the number of positives out of the total studies. The numbers refer to reliable studies, identified according to the criteria explained above.

Table 3. Total number of reliable genotoxicity studies for nano-TiO2 found in the literature: number of positives / number of total studies. The comet *in vitro* assay is the most performed assay and it gives mostly positive results for NM-100, NM-102 and NM-105.

	Name	NM-100	NM-101	NM-102	NM-103	NM-104	NM-105
In vitro	Micronucleus assay	-	-	3/10	3/8	3/8	4/18
	Comet	2/2	2/6	5/8	0/6	0/6	10/14
	Chromosomal aberration	-	-	-	-	-	0/1
In vivo	Micronucleus assay	-	0/3	0/6	0/5	0/5	2/9
	Comet	-	1/5	2/13	1/12	2/12	4/15
	Chromosomal aberration	-	-	0/2	-	-	-

3.3. Data matrix

The ECHA guidance on grouping for read-across (ECHA 2017a) suggests to build a matrix (fourth step of the proposed framework) to report the data collected and evaluated in the previous step.

Table 5 includes the physicochemical properties of the source and target NMs ("what they are"). For the analogues information on fundamental behaviour ("where they go") and reactivity ("what they do") is reported, as well as the toxicological information, i.e. the results of the *in vitro* comet assay.

4. HYPOTHESIS FOR THE CATEGORIES

Based on the data collected an initial grouping of the identified materials is proposed. The development of the category hypothesis (indicated as second step of the ECHA guidance for grouping of nanoforms, ECHA 2017a) was an iterative process in the case study. It is described here after the data collection.

4.1. Identification of the nanoforms within each group

According to the trend identified for the *in vitro* comet results (see Table 3), the test has positive results (the group of positives) for NM-100, NM-102 and NM-105. The group of materials that are not causing DNA damage in the *in vitro* comet assay (the group of negatives) is NM-101, NM-103 and NM-104. This grouping is schematically presented in Table 4.

Table 4. Assignment of the analogues to the two groups identified regarding DNA damage potential measured with the *in vitro* comet assay, depending on the NM characteristics.

	Naked particle	Particle with high non-TiO_2 content (coating and impurities)
NMs	NM-100, NM-102, NM-105	NM-103, NM-101, NM-104
DNA damage	+	-

4.2. Development of the initial grouping hypothesis

The dataset includes six NMs with different properties: different primary and crystallite sizes (5-100nm), different crystalline types (anatase and rutile) and surface characteristics (coated or uncoated). From combining the *in vitro* comet assay results and the physicochemical characteristics of the nanoforms, as shown in Table 5, the following grouping hypothesis can be formulated:

If nano-TiO_2 is uncoated, it has the potential to damage DNA. If a non-reactive coating and large amounts of impurities are present on the surface of the NM, this can mask the effect and no DNA damage is caused.

The hypothesis is based on the fact that the data show that NM-100, NM-102 and NM-105 have a higher tendency to giving positive results in the *in vitro* comet assay, whereas NM-101, NM-103 and NM-104 tend to give negative results (see Table 3).

Table 5. . Data matrix showing the six analogues and two target nanomaterials: physicochemical properties and selected toxicological information.

		SOURCE ANALOGUES						TARGET NMS	
	Name	NM-100	NM-101	NM-102	NM-103	NM-104	NM-105	TiO_2 R	TiO_2 A
TOXICOLOGICAL INFORMATION	*In vitro* comet assay	+	-	+	-	-	+	?	?
PHYSICOCHEMICAL PROPERTIES	Total non-TiO_2 content including coating and impurities (% w/w)	1.5	9	5	11	11	0.11	13	0.5
	Surface coating (%)	0	0	0	8	8	0	11	0
	Surface chemistry (as declared by manufacturer)	uncoated	uncoated	uncoated	Al_2O_3 and SiO_2	Al_2O_3 and SiO_2	uncoated	SiO_2, Na_2SO_4, SO_4	uncoated
	Organic matter (%)	0	8	0	2	2	0	9	0
	Impurity(% w/w Fe)	0.49	0	0.07	0.06	0	0.06		
	Impurity(% w/w Si)	0.28	0.29	0.08	0.68	0.018	0.07		
	Impurity(% w/w K)	0.25	0	0.001	0.001	0.001	0		
	Impurity(% w/w P)	0.21	0.27	0.001	0	0	0		
	Impurity – coating (% w/w Al)	0.09	0.09	0.05	3.4	3.2	0.04		
	Impurity(% w/w Cr)	0.03	0	0	0	0	0		
	Impurity(% w/w Zr)	0.005	0.01	0.005	0.001	0.001	0		
	Impurity(% w/w Ca)	0.001	0	0.005	0.005	0.01	0		
	Impurity(% w/w Na)	0.001	0.1	0.001	0.01	0	0.001		
	Impurity(% w/w S)	0	0.22	0.001	0.01	0.01	0.26		
	Impurity(% w/w Mg)	0	0	0	0.001	0.001	0		
	Crystal type (Anatase)	1	1	1	0	0	0.84	0	1
	Crystal type (Rutile)	0	0	0	1	1	0.16	1	0
	Crystal type (Cubic)	0	0	0	0	0	0	0	0
	Crystallite size (mean)	117.81	7.69	23.93	24.32	24.71	22.44		
	Primary particle diameter (mean)	93.45	5.25	22.00	24.00	24.50	20.13	62x10	14
	Shape (elongated/rod=1, spherical=0)	0	0	0	1	0	1	1	0
	Aspect ratio	1.53	1.53	1.53	1.70	1.53	1.36	6.2	1
	Specific surface area (m^2/g)	9.23	316.07	77.86	53.98	54.33	47.00	149	177
	Particle size distribution (nm)	210	278	440	135	145	177		
	Particle size distribution in PBS, untreated (nm)	2289	1229	1579	1397	1600	3342		
	Particle size distribution in MQ, M1, 1 min (nm)	259.3	719.5	703.0	2649.0	207.7	352.6		

	Particle size distribution in MQ, 1min (nm)	201.3	500.9	505.7	1977.0	194.3	227.5	
	Zeta Potential in MQ, 1min (mV)	-24.5	-27.2	-27.1	39.1	-23.4	-23.8	
	Zeta Potential DMEM 10 FBS, M1, 1 min (mV)	78.40	0.13	-10.50	-12.40	-9.38	-9.92	
	Zeta Potential PBS, 20 min (mV)	-20.2	-21.7	-18.5	-20.9	-20.3	-33.2	
	Zeta Potential DMEM 5 FBS, 20 min (mV)	-10.4	-11.3	-9.5	-13.7	-9.4	-11.9	
	PdI in PBS, untreated	0.219	0.239	0.769	0.255	0.232	0.232	
	PdI in MQ, 1 min	0.205	0.274	0.248	0.393	0.236	0.211	
	PdI in DMEM & L-glutamine, 20 min	0.515	0.247	0.227	0.264	0.209	0.341	
	PdI	0.303	0.323	0.427	0.292	0.227	0.245	
Where they go	Isoelectric Point (Mean) (pH)	NA	5.5	6.0	8.3	8.5	6.8	
	Isoelectric Point (Min) (pH)	NA	5.3	6.0	8.2	8.2	6.6	
	Isoelectric Point (Max) (pH)	NA	5.7	6	8.5	8.8	6.9	
	Density (g/mL)	3.84	3.99	3.84	4.02	4.09	4.05	
	Mean of total pore volume (mL/g)	0.032	0.319	0.300	0.262	0.194	0.194	
	Micropore surface area (m^2/g)	0	13.625	1.108	0	0	0	
	Micropore volume (mL/g)	0	0.00179	0.00034	0	0	0	
	Dustiness-Respirable(mg/kg)	1500	5600	9200	19000	6400	11000	
	Biodurability 24h 0.05% BSA (Ti content) (µg/l)	5.2	0	0	0	0	0	
	Biodurability 24h Gambles solution (Ti content) (µg/l)	0	0	3388	0	0	0	
	Biodurability 24h Caco2 (Ti content) (µg/l)	796	3414	1741	222	3386	2724	
What they do	Redox caco2 medium [a]	1	-1	-1	1	-1	-1	
	Redox Gamble's solution [a]	1	0	-1	1	-1	-1	
	Redox BSA [a]	0	0	0	0	0	0	

a values obtained from Nanogenotox Deliverable 4.7 determined by measuring the content of O_2. Oxidising properties (1), neutral (0), reducing (-1)
MQ: milli-Q water; 1 min: sonication time, as direct probe; 20 min: sonication time, ultrasound batch mode; DMEM: Dulbecco's modified eagle medium; PBS: phosphate-buffered saline (PBS: 137 mM NaCl, 2.7 mM KCl, 0.02 M PO4) FBS: fetal bovine serum; PdI: polydispersity index; BSA: bovine serum albumin.

As also reported in the literature (Magdolenova et al. 2014; Golbamaki et al. 2015), there are more experimental data on the *in vitro* comet assay for testing the DNA damage caused by NMs, compared to other tests. The *in vitro* comet assay detects DNA strand breaks at the level of single cells. In the case study, the aim was to identify NMs physicochemical properties that may affect DNA damage, to be able to read-across test results to the target NMs.

Regarding the mechanism(s) of genotoxic action of NMs, Magdolenova et al. (2014) and Golbamaki et al. (2015) reviewed the mechanisms of genotoxicity of NMs and of the subset of metal oxides (including nano-TiO_2), respectively. They report that DNA damage caused by nano-TiO_2 may be classified as direct primary genotoxicity, indirect primary damage, and secondary genotoxicity. Direct genotoxicity assumes that DNA and NM are in contact. An example of direct genotoxicity by TiO_2 was found to involve the reaction of terminal DNA phosphate groups that influence the binding of DNA to nano-TiO_2 (Rice et al. 2009). Indirect primary genotoxicity may be elicited by interaction of NMs with nuclear proteins (involved in replication, transcription, and repair), disturbance of cell cycle checkpoint functions, reactive oxygen species (ROS) arising from the NM surface, release of toxic metal ions from the NM surface, ROS produced by cell components, and inhibition of antioxidant defence (Jugan et al. 2012). Finally, secondary genotoxicity may be elicited by ROS production in inflammatory cells via an inflammation signalling pathway (Romoser 2012), i.e. macrophages or neutrophils (activated phagocytes) generate ROS while trying to digest the NMs. This can cause an inflammatory reaction that may subsequently cause oxidative DNA damage (Trouiller et al. 2009).

Although most experimental studies provide evidence for a mechanism of action for indirect primary genotoxicity via ROS (Golbamaki et al. 2015), several studies report that a clear correlation between the level of ROS production and DNA damage was not supported by their findings (Barillet et al. 2010, Golbamaki et al. 2015, Li et al 2013). For example, Li et al. (2013) investigated the ability of a set of NMs (including nano-TiO_2, plus metal and metal oxide NMs and quantum dots) to inhibit DNA replication by binding to DNA: they concluded that ROS generation as an important cause of genotoxicity was not supported by their experiments as the amount of generated ROS did not explain the effect of NMs on DNA binding reported in their study. They suggested instead that direct binding activity of NMs to DNA was the likely genotoxicity mechanism.

It can be readily seen in the dataset of analogues (matrix Table 5) that the coated NMs turn out negative in the comet assay while the ones without coating and organic impurities turn out positive. This can be explained if a direct interaction mechanism of genotoxicity or an indirect primary genotoxicity are considered (Magdolenova et al. 2014). The conduction band of TiO_2 falls in the range of biological redox potentials (Burello & Worth 2011), meaning that TiO_2 with or without the presence of UV light can generate reactive species that react with cell constituents such as DNA. In both direct and indirect primary genotoxicity, physical interaction of the NM with DNA (direct) or another cellular component (e.g. enzyme mediated a redox reaction) that generates ROS (indirect) is necessary for the DNA damage to occur. The NM coating acts as a physical barrier that can prevent this contact between the Ti and O atoms of TiO_2 and DNA or other cellular components. Therefore, coated nano-TiO_2 will not turn out positive in the comet assay as there will be no physical interaction between the Ti / O atoms and DNA / cellular components.

The way in which the coating can prevent DNA damage is not entirely clear, in fact several works show contradictory results and explanations for the i*n vitro* genotoxicity of TiO_2 with coating playing a main role. For instance, it was shown (Mano et al. 2012; Falck et al. 2009) that the addition of PEG coating to nano-TiO_2 increased the dispersion of NMs which resulted in lower cytotoxicity and genotoxicity. Magdolenova et al. (2012) showed that the degree of dispersion of TiO_2 NMs had an influence on the DNA damage in three cell lines. Agglomerates of less than 200 nm had no effect on genotoxicity while larger ones showed positive results. These results could be due to larger agglomerates that precipitate and deposit on the cells increasing the actual exposure to the NM or even covering them completely and suffocating them. Another consideration is the effect that the use of media with proteins (e.g. BSA, FBS) can have on the results. If the NMs are surrounded by proteins, they are more dispersed and also less toxic as the "reactive" part is encapsulated ("hidden") behind the protein corona. Another aspect that cannot be ignored when analysing the *in vitro* results of TiO_2 is its photocatalytic activity, which can be even triggered by a simple fluorescent tube (Karlsson et al. 2015). Thus, it is obvious that the mechanism of genotoxicity of TiO_2 is not well defined and that there might be more than one that could even take place simultaneously. Probably the truth is a combination of all factors that have as common source the presence of coating either by preventing aggregation of NMs, deposition, and therefore reducing exposure, or by preventing physical contact with DNA and/or other cell components after uptake. However, what is relevant in this case is that the majority of studies agree with the hypothesis presented here which is the fact that coated nano-TiO_2 show fewer positive results in the comet assay than the uncoated ones, therefore it can be fairly concluded that the presence of coating reduces the number of positive *in vitro* comet assay results for nano-TiO_2. It is important to keep in mind that the present coatings (Al_2O_3 and SiO_2) are mainly not "charged" as could be coatings with –COOH or $–NH_2$ terminal groups, and that their conduction band falls outside the biological redox band (Burello and Worth, 2011). Different coatings may result in a different grouping hypothesis.

5. JUSTIFICATION OF DATA GAP FILLING

The fifth step of the ECHA draft guidance on grouping for read-across (ECHA 2017a) aims at combining all the gathered information into an overall assessment and filling the data gaps.

5.1. Assessment of the similarity hypothesis for the categories

The identification of the similarities[2] in the analogues dataset is a way to assess the strength of the stated grouping hypothesis. Given that there is not an agreement in the literature about the mechanism that can lead to genotoxicity of nano-TiO_2, e.g. chemical reactivity, ROS generation, agglomeration and sedimentation, a set of statistical methods often used in chemoinformatics were applied using R 3.2.5 (R CoreTeam, 2016) to identify the (physicochemical) properties that differentiate the analogues, determine their similarity and that may drive genotoxicity. Thus, the grouping hypothesis was based on the similarity between the physicochemical properties reported in the dataset. The formulation of similarity rules between nanoforms was supported by the application of the following chemoinformatics data mining methods:

1. Hierarchical clustering (Suzuki et al. 2014) to identify possible clusters or groups of analogues in the dataset (similar NMs)
2. Principal component analysis (Husson et al 2006) for an indication of the physicochemical properties that differentiate the NMs, justifying the clustering of the grouping hypothesis
3. Variable selection with random forest (Liaw et al 2002) to verify which are the most relevant properties in predicting the assay results. For this methodology the toxicological information needs to be taken into consideration as the aim was to classify the relevance of each variable (physicochemical property) in predicting the result (toxicological assay). In this case the toxicological endpoint was genotoxicity but the same technique could be used for any other endpoint, the result being exclusive to each endpoint.

Hierarchical clustering

The initial dataset included six analogues with approximately 147 properties for each of them. Four properties were discarded because of low variability. If a correlation filter was applied to such a small dataset, the correlation would be derived from just six values and this could lead to the filtering of properties that are not really related. For instance, the correlation between "Organic matter" and "Micro surface area" is 0.93. Such a correlation is not meaningful because the former property only contains values for three

[2] In this context similarity includes the similarity in the properties under the nanoforms identification, fundamental behaviour and reactivity

NMs, and the latter for two. Thus, a correlation filter was not applied in order to avoid such problems.

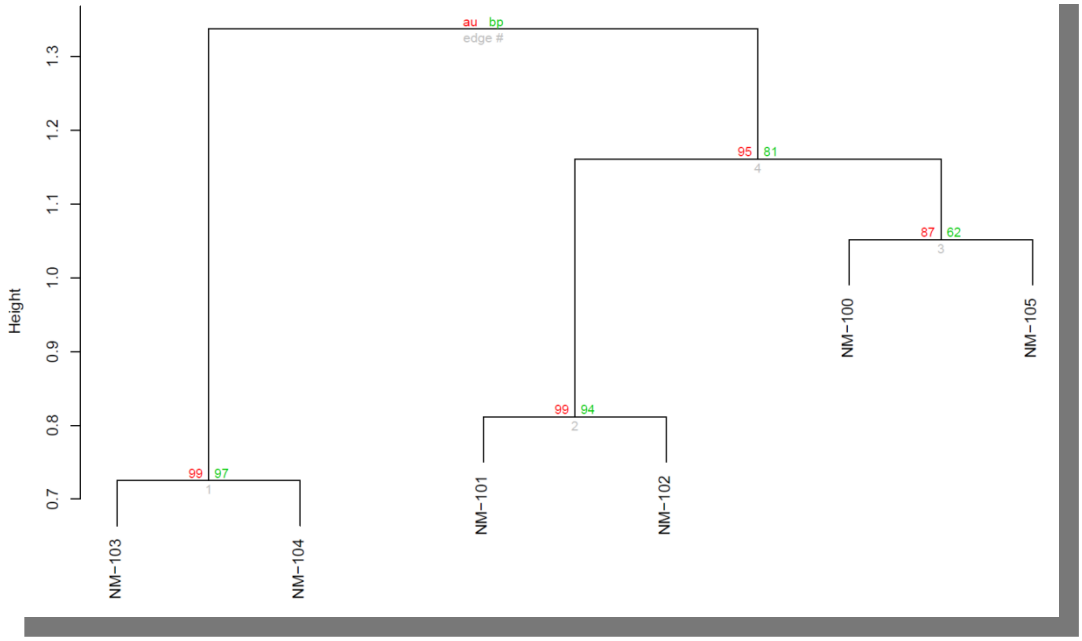

Figure 2. Hierarchical clustering of the TiO_2 analogues with representative particle size distribution, Zeta potential and polydispersity index. The numbers in red correspond to the "Approximately Unbiased" (AU) p-value that is computed by multiscale bootstrap resampling, and the ones in green to "Bootstrap Probability" p-value (BP), which is computed by normal bootstrap resampling. The height in the Y-axis indicates the distance between clusters computed as average linkage. The AU p-value will be used for the interpretation as it is usually a better approximation to the real p-value.

The dataset contained a number of dynamic light scattering DLS measures for the NMs in 6 different solvents (MQ water, PBS, DMEM + 1% FBS, DMEM + 5% FBS, DMEM + 10% FBS, and DMEM + L-glutamate) and 3 different treatments (untreated, 1min tip sonication, 20min bath sonication). This made a highly biased dataset as it contained 62 properties related to DLS particle size distribution, 21 to Zeta potential, and 20 to polydispersity index (PdI). In order to reduce the weight of such measures and obtain a more balanced dataset, these properties were reduced to 4 DLS measures, 4 Zeta Potential and 4 PdI. A hierarchical clustering of the transposed dataset was used to determine clusters of similar properties (e.g. particle size distribution in DMEM + 5% FBS and 1min sonication with particle size distribution in DMEM + 10% FBS and 1min sonication). This allowed the determination of groups of similar properties which could be reduced to a single property and yield a more balanced dataset. Four groups were considered for each type of property, i.e. particle size distribution, Zeta Potential, and PdI, and one property was kept for each set of related properties and the others were discarded.

The hierarchical clustering of the resulting dataset, which contained 50 variables, is presented in Figure 2 and shows that NM-103 and NM-104 form a very solid group ($p<0.01$). The other 4 four NMs form another group since, although NM-101 and NM-102 are clustered together with high significance ("Approximately Unbiased" (AU)

value), NM-100 and NM-105 do not show such a high strength, and therefore the higher level group prevails ($p<0.05$). It is worth mentioning that the clusters obtained here must be only considered from an exploratory point of view and in a weight of evidence context. This information alone cannot be used to define clusters of NMs but must be complemented with other techniques and rationales (e.g. PCA, variable selection, mechanistic information) to be used in read-across.

Principal component analysis

While the hierarchical clustering indicates similar NMs by taking into account all physicochemical properties and forming subsequent groups of two substances, the principal component analysis (PCA) is a dimensionality reduction technique that shows the properties that account for the maximum variance between individuals, NMs in this case. The PCA also uses all properties to determine each of the principal components (PC) but they are weighted in such a way that a minimum number of properties can be used to explain the differences between the different NMs.

The PCA of the same dataset used for the hierarchical clustering of the analogues shows a similar picture (Figure 3) to the one obtained in the hierarchical clustering. The different NMs are placed in the plot by using the PC1 and PC2 scores and the loadings of each property with respect to PC1 and PC2 are indicated as arrows. NMs that appear close to each other indicate similarity in the space defined by PC1 and PC2. Long and light blue arrows indicate high contribution of that specific property to one of the PCs. The closer the arrow is to an axis, i.e. to a PC, the higher the contribution it has to that PC. It is necessary to remember that the PCs are a simplifications of the whole picture and that the fact that NMs appear close to each other only indicates that these NMs are similar to each other in that reduced representation of reality given by two variables, usually PC1 vs PC2. PC1 and PC2 typically account for a rather large variance (>50%) – PC2 is 25.5% and PC1 is 36.2% ─ and indicate what are the variables that differentiate the NMs. The fact that these variables be related with the endpoint of interest cannot be assured and is not the purpose of PCA or other unsupervised techniques.

In Figure 3 NM-103 and NM-104 appear close to each other at the positive side of PC2. The arrows show that these positions are mainly driven by the properties related to impurities of Al, Mg, by the crystal type rutile, and % of surface coating. NM-100 appears at the top part of the plot, mainly driven by particle primary diameter and crystallite size as it is the biggest NM of the series (considered as bulk material). For the same reason, NM-101 appears at the bottom of the plot as it is the smallest NM, and NM-102 and NM-105 appear next to each other on the negative side of PC1, mainly driven by crystal type anatase and no surface coating.

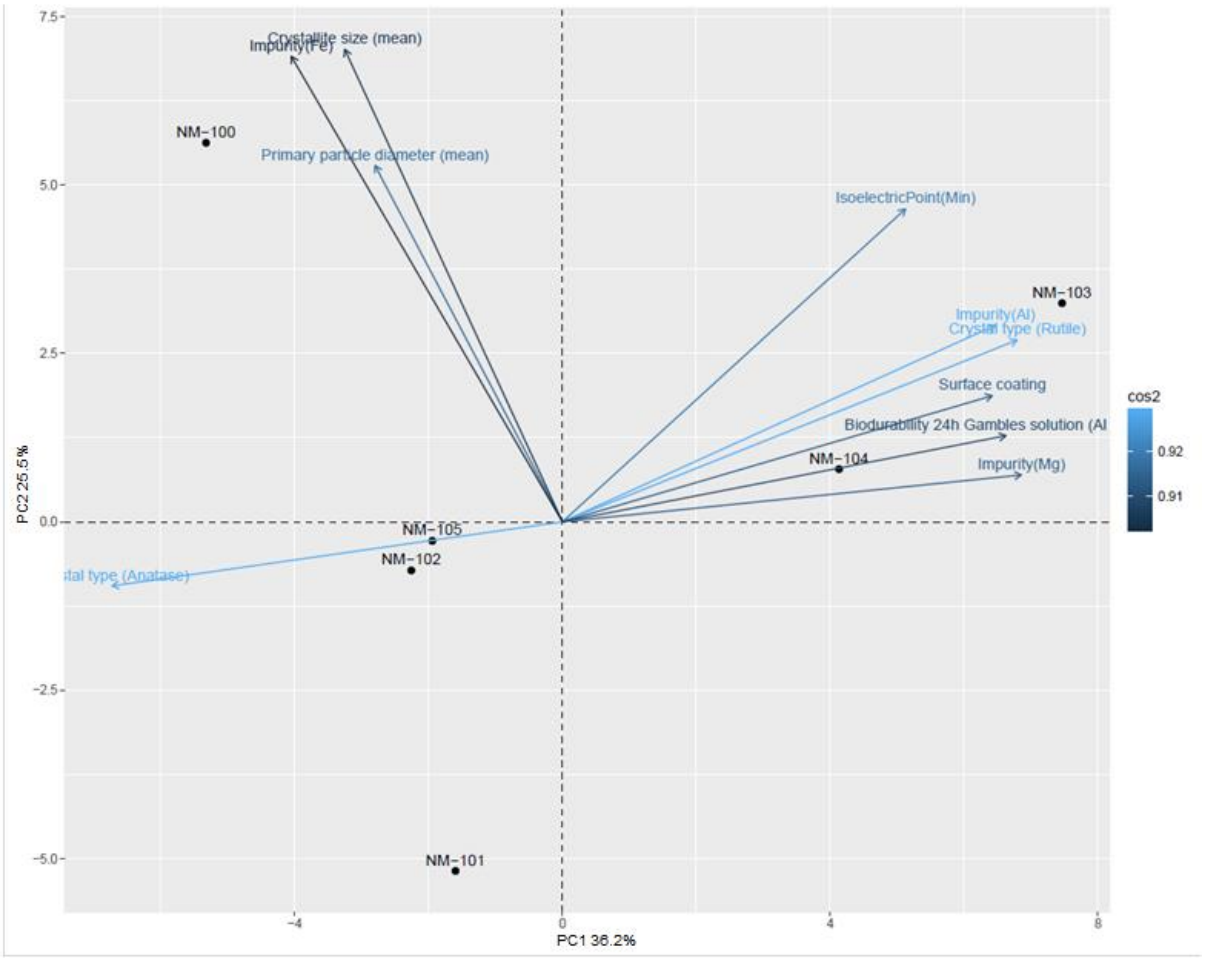

Figure 3. Principal component analysis of the dataset of six TiO_2 analogues. The position of the analogues (individuals) on the space of PC1vsPC2 are indicated as black dots. Arrows correspond to the 10 variables with higher contribution to the principal components (PCs). The colours are defined by the squared loadings (cos2) and indicate their contributions to the PCs.

The squared loadings of the two first principal components are given in Table 6 and show that the properties with the higher contributions to PC1 are crystal type (anatase and rutile), biodurability and Al impurity, which are similar properties, and % of surface coating and Mg impurity. For PC2 the main contributors are specific surface area, total pore volume, primary particle diameter, crystallite size and Fe impurities.

It can be deduced from the PCA that the differences between NMs are mainly due to particle type (anatase vs rutile), particle size, and surface coating. The fact that crystal type variables appear so high in the list is due to the fact they correspond to the percentage of crystal of that type, and since most of the particles are either 100% anatase or 100% rutile, the differences between them are extreme. Particle size diameter is also one of the main differences between NMs as the biggest one is 115 nm and the smallest is 5 nm. Biodurability(Al) and Impurity(Al) are very similar properties as the former one corresponds to the quantity of Al dissolved in media after 24h, and the second one corresponds to the quantity of Al found after calcination of the NMs. The latter, is also part of coating as NM-103 and NM-104 have around 6% of Al_2O_3 coating and therefore

a large amount of Al. Surface coating, which includes not only Al_2O_3 but also organic matter like silanes or glycol, is also one of the main contributors to PC1. The loadings also show that other properties like Zeta potential, PdI, or particle size distribution have less influence.

Table 6. Squared loadings of PC1 and PC2 of the PCA of the analogues.

Property	*PC1 loadings²*	*Property*	*PC2 loadings²*
Biodurability 24h Gambles solution (Al content)	0.90	Specific surface area (mean)	0.77
Impurity(Al)	0.89	Mean of total pore volume (ml/g)	0.74
Crystal type (Rutile)	0.89	Primary particle diameter (mean)	0.73
Crystal type (Anatase)	0.89	Crystallite size (mean)	0.67
Surface coating	0.87	Micropore volume (ml/g)	0.63
Impurity(Mg)	0.87	Impurity (Fe)	0.63

Random forest as variable selection to predict the comet assay

Hierarchical clustering and PCA make use of the physicochemical properties of substances to determine clusters of similar substances and the properties that differentiate those substances the most. The random forest variable selection algorithm, on the other hand, uses the physicochemical properties to predict a given outcome, in this case positive or negative results in *in vitro* comet assays, and provides a relative importance of the variables for the prediction.

The variable importance plot of the analogues (Figure 4) clearly shows that the most important variables to predict the comet assay results for the six analogues are the amount of organic matter and Total non-TiO_2 content. As mentioned above, Total non-TiO2 content corresponds to the sum of the impurities and coating found on the analogues, and this distinction had to be made because it is impossible to determine from the data whether NM-101 was coated. According to the manufacturer, NM-101 has no coating, however the tests performed in the Nanogenotox Joint Action showed that there was a significant mass loss at 200 °C that was identified by gas chromatography – mass spectrometry (GC-MS) analysis as hexa/octadecanoic acid and others. These impurities accounted for around 9% of the total weight and were considered as coating by Nanogenotox. They were not considered coating in this case study because NM-101 is stated to be uncoated by the manufacturer, and according to REACH, only those materials can be considered coating that were intentionally added. Therefore, the 9% of organic matter of NM-101 has been considered as “impurities” here.

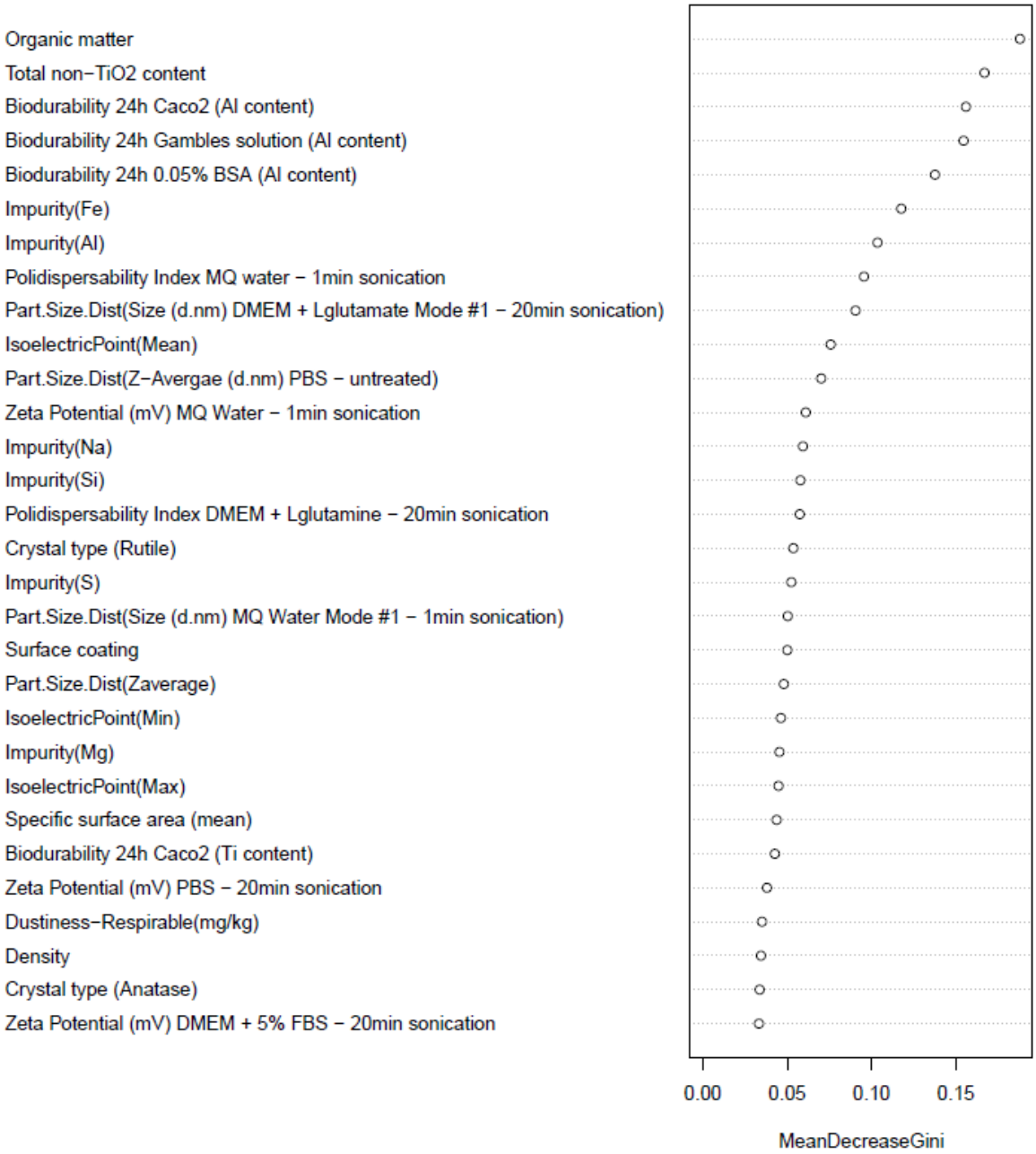

Figure 4. Relative importance of variables in terms of their predictivity of the *in vitro* comet assay. Variable importance expressed as mean decrease of the Gini index of the analogues.

Conclusion from chemoinformatics similarity analysis

Hierarchical clustering and PCA of the analogues show that two groups of NMs can be clearly defined from their physicochemical properties. NM-103 and NM-104 (negative in the *in vitro* comet assay) form a very strong group ($p<0.01$). Actually, they are very similar NMs of rutile type with a size of ~24nm, high content of impurities like Al, and with surface coating. The coating in fact contains Al_2O_3, which explains that Al

impurities content was found as a relevant property. Another group is the one formed by NM-102 and NM-105 (positive in the *in vitro* comet assay) as both correspond to uncoated anatase TiO_2 (100% and 84%, respectively) with ~23nm and a low amount of impurities. NM-100 does not cluster together with any of the other NMs in the PCA. The reason for this is that it corresponds to a relatively large "NM" (>117nm), which makes it significantly different from the rest. For instance in the PCA, PC2 has a strong component of particle size and, therefore these property sets NM-100 at the higher part of the plot. However, if only the crystal type and coating are considered, NM-100 groups perfectly with NM-102 and NM-105 as it is uncoated, and 100% anatase. In addition, such a classification matches the toxicological profile of these NMs as they all turn out positive in the *in vitro* comet assay. NM-101 is more difficult to classify because it is the smallest of all NMs with a diameter of 5nm (lower part of PCA), and it is also declared as uncoated by the producer, whereas the elemental analysis showed the presence of around 9% of organic matter or impurities. Moreover NM-101 is of type anatase but negative in the *in vitro* comet assay.

If NM-101 is not considered, it can be clearly stated that uncoated anatase NMs are the ones positive in the *in vitro* comet assay. This in fact is in line with the higher photoreactivity of anatase with respect to rutile (Luttrell et al. 2014). Although comet assays are carried out in lab facilities where very little UV light should be present to differentially photoactivate anatase, it was recently shown that anatase samples exposed to fluorescent tube lab light caused statistically significant higher amounts of DNA breaks (Karlsson et al. 2015) to BEAS-2B cells. However, NM-101 is also of anatase type but turns out negative in the *in vitro* comet assay, thus breaking the abovementioned relation between anatase and positive comet assays. NM-101 contains a large amount of organic impurities (9%), which is of similar composition to the coating of NM-103 and NM-104. Whether these organic impurities (organic matter in Table 5) are considered coating or not – in order to be considered coating the substance must be intentionally added – it is clear that there is a correlation between the NMs that have coating and/or organic impurities and the result of the *in vitro* comet assay. Moreover, this correlation is supported by the results of the random forest variable selection which shows that the two variables with the highest discriminant power to predict the *in vitro* comet assay are Total non-TiO_2 and Organic matter. Following the hypothesis, uncoated rutile would also be predicted as possibly genotoxic. However, it would be desirable to dispose of data for this type of nanoform in order to be able to have a prediction with less uncertainty.

5.2. Filling data gaps by read-across

The grouping hypothesis for the nano-TiO_2 analogues was supported by the results obtained from hierarchical clustering, PCA and random forest variable selection algorithms.

The two target NMs were identified as described in Table 1, including the coating of the two nanoforms. According to the physicochemical properties of the identified target NMs, it can be assumed that they are included in the same variable space as the analogues: primary particle size, shape, total non-TiO_2 content, organic matter, crystal type, and specific surface area are included in the range defined by the analogues. Because of the lack of some physicochemical data for the target NMs, it was not possible to include them in the PCA analysis or in the clustering exercise. However, it is possible to assign the two target NMs to a class according to some of their characteristics. Since

the presence of coating or high amount of non-TiO_2 content on the surface of nano-TiO_2 appears to prevent NM to cause DNA damage detected by the *in vitro* comet assay, it is possible to group TiO_2 R nano with the analogues NM-103 and NM-104 and possibly NM-101, giving negative results, and TiO_2 A nano with NM-100, -102 and -105, which cause DNA damage.

As shown in Table 5, the two target NMs have different characteristics with respect to coating or non-TiO_2 content. TiO_2 R has a coating, and thus it is predicted to have a negative outcome in the *in vitro* comet assay. TiO_2 A, on the contrary, has a relatively low level of impurities and no coating, and thus the prediction is that it gives a positive result in the *in vitro* comet assay. The prediction is reported in Table 7.

This outcome is confirmed by the *in vitro* comet assay carried out by Guichard et al. (2012) which shows that TiO_2 A is genotoxic and TiO_2 R is not.

Table 7. Data matrix showing the six analogues and the two target nanomaterials with read-across prediction results.

Name	NM-100	NM-101	NM-102	NM-103	NM-104	NM-105	TiO_2 R	TiO_2 A
In vitro comet assay	+	-	+	-	-	+	-	+

Perform and/or propose testing

The sixth step of ECHA draft guidance (ECHA 2017a) asks for identifying testing, where the data collected to support the grouping hypothesis is not considered reliable or sufficient.

The hypothesis is based on a small dataset and considers a single *in vitro* endpoint. Other *in vitro* test results available for nano-TiO_2 do not confirm genotoxicity, so an overall conclusion on genotoxic potential cannot be made with certainty. Extension of the dataset with reliable *in vivo* genotoxicity data could be a possible next step. In *in vivo* studies, toxicokinetics data show that after intravenous administration, liver, spleen and lungs may be considered as target organs as there is evidence that nano-TiO_2 reaches these organs and tend to accumulate there. Since biological reactivity is an important property in supporting similarity, it would be interesting to have a test guideline to fulfil to this requirement, as at the moment there are a few data available and there is no standardised experimental approach to investigate this property.

Following the hypothesis, uncoated rutile would also be predicted as possibly genotoxic. It would be desirable to obtain data for this type of nanoform (uncoated rutile; as well as intentionally coated anatase) in order to strengthen the prediction.

6. CONCLUSION ON THE READ-ACROSS PREDICTION

With support of cheminformatics tools and analysis of the collected data for the analogues, a link between nanomaterial properties and the outcome of the *in vitro* comet assay could be established, following the hypothesis that nano-TiO_2 in its uncoated form has the potential to damage DNA, which can be masked by the presence of non-reactive coating and large amounts of impurities on the NM surface. The prediction of a negative result in the *in vitro* comet assay for the coated TiO_2 R and of a positive result for the uncoated TiO_2 A is confirmed by experimental data.

7. UNCERTAINTY ASSESSMENT

Different areas contribute to the overall uncertainty of the read-across process: uncertainty associated with the data, the similarity hypothesis to define analogues, and the read-across prediction of hazard derived (see Schult et al. 2015).

For the present case study the ECHA Read-Across Assessment Framework (RAAF) (ECHA 2017b) was used as a systematic guidance to identify and summarise the different sources of uncertainty associated with filling data gaps by grouping and read-across, with a view also to evaluate the applicability of the RAAF for nanomaterials and to find possible issues that need to be taken into account particularly for nanomaterials. Generally the reliability, relevance, consistency and completeness of data and argumentation are evaluated.

7.1. Evaluation of the read-across according to the Read-Across Assessment Framework

The Read-Across Assessment Framework has been developed by ECHA to provide guidance for a structured analysis of read-across submissions and justifications (ECHA 2017b). Six possible read-across scenarios are considered, depending on the number of substances (analogue or category approach), the effect caused by common or different substances for source(s) and target(s) or, for a category, whether the predicted property is following a regular pattern (trend) or not changing across source structures. Sets of Assessment Elements (AEs) per scenario describe 'crucial scientific aspects to judge validity and reliability of read-across' (ECHA 2017b). For each AE, multiple considerations are listed, which should be addressed in the justification of the read-across argumentation. There are general (common) AEs and scenario-specific AEs. Assessment Options (AOs) reflect the conclusions on adequacy and scientific robustness. They are defined as scores from 1 to 5 according to whether the information provided is not acceptable (1), not acceptable in its current form (2), acceptable with just sufficient (3), medium (4) or high (5) confidence.

From the six RAAF scenarios, scenario 6 describes the TiO_2 nanomaterial case study best. It uses a category approach (reading-across from a group of substances to the targets), different compounds (in this case nanoforms) are considered in the category to cause the same type of effect, and there is no regular trend of an effect (potency) associated to the source structure. (The comet assay result is binary, either positive or negative.)

The AEs for RAAF scenario 6 cover the general overarching topics:

- similarity hypothesis, substance(s) considered, and available data
- toxicant(s) the organism is exposed to
- mechanism of toxicity.

They are listed in Table 8. AEs C.1-C.6 are common AEs for all scenarios considering a category approach; AEs 6.1-6.5 are scenario-specific. The terms used correspond to

the terminology used in the RAAF, i.e. applicable to conventional organic substances. Nanospecific adaptions of the terms will also be discussed in the following.

The RAAF AE elements (scenario 6) are used to highlight uncertainties in the grouping and read-across argumentations in this case study. The evaluation and nanospecific considerations are summarised in Table 8.

Table 8. Evaluation of the uncertainties of the TiO_2 case study according to the ECHA RAAF scenario 6.

	RAAF Assessment Element (Scenario 6)	Uncertainties in the TiO_2 case study	Nanospecific issues
C.1	Substance characterisation	• Measured physicochemical characteristics of the NMs vary: measurement uncertainty. Is there an influence on other properties of the nanomaterials? • Impurity information not always available or inconsistent	• Physicochemical characterisation of NMs: high variability of measurements (influence of different experimental conditions)
C.2	Structural similarity and category hypothesis	• NM-101 is not declared as coated, but has 9% organic impurities corresponding to a coating. Thus it was considered coated. • Different composition of the coatings/impurities (e.g. some containing Al_2O_3) • Possible influence of crystal type if particle not coated. • Uncertainty of reading across a spherical to a rod-shaped particle.	• For NMs, the similarity cannot be based on chemical (e.g. molecular) structure as for conventional chemicals, but should consider physical form and key physicochemical properties
C.3	Link of structural similarities and structural differences with the proposed property	• Little is known about the mechanisms of toxic action, making it challenging to link similarity to the property (genotoxicity) considered	
C.4	Consistency of effects in the data matrix	• Uncertainty in applying existing testing protocols to nanomaterials and thus uncertainty in assessment of quality, reliability and relevance to human health endpoints of measured toxicity data	• Artefacts affecting the results of toxicity assessment of NMs are discussed in the literature
C.5	Reliability and adequacy of the source study(ies)		
C.6	Bias that influences the prediction	• Selection of analogues based only on data availability	
6.1	Compounds the test organism is exposed to	• The mechanism of genotoxicity of TiO_2 is not well defined. It is also possible that several effects take place at the same time.	• For conventional chemicals, either the parent molecule or (bio)transformation products are the indirect/direct toxicants; for NMs the considerations extend to coating, released metals etc
6.2	Common underlying mechanism, qualitative aspects		
6.3	Common underlying mechanism, quantitative aspects		
6.5	Occurrence of other effects than covered by the hypothesis and justification		
6.4	Exposure to other compounds than to those linked to the prediction	• For example the presence of reactive transition metals may also contribute to oxidative DNA damage induction.	

Substance characterisation (C.1)

In the first step of the read-across process is the identification of the nanoforms considered, i.e. “what they are”. The identification includes particle size, particle shape and surface chemistry. The characterisation should also include impurities present, which are defined as "unintended constituents present in a substance as manufactured" (ECHA 2017d); on the other hand, NMs contain also surface coating purposely added.

The core chemical composition is TiO_2, the crystal type and size are reported. Impurity information is given by the provider only as a generic percentage of purity. Additional measurements are available from different Nanogenotox deliverables; however, in some cases there were some inconsistencies between different tests.

For the considered target substances, analysis of the physicochemical properties (as reported in Guichard et al. 2012) showed that the measured properties were slightly different from those reported by the manufacturer.

Of the physicochemical properties collected, some were disregarded because of lack of information or low data reliability. The analysis of the variability of the physiochemical property measurements has shown particle size distribution and zeta potential being dependent for example from the dispersion medium used or different sonication methods. As a consequence of the huge data variability in particle size distribution (and consequently on zeta potential), this measurement is not reliable and could in principle be excluded from the grouping exercise.

Overall there is uncertainty associated with the nanoform identification and physicochemical characterisation, which is subject to high variability in measurements (different experimental conditions, result ranges).

Structural similarity and category hypothesis (C.2)

The criterion of structural similarity considered in the RAAF for conventional organic substances needs to be extended for nanoforms to their properties regarding identification, fundamental behaviour and reactivity.

In order to investigate criteria for similarity of the nanoforms considered and form a category hypothesis, chemoinfomatics methods have been used to identify the (physicochemical) properties that differentiate the analogues, determine their similarity and that may drive genotoxicity (as detected by the *in vitro* comet assay). The properties differentiating the two groups were surface coating, Al coating and crystallite type.

The nanoform NM-101 was not declared coated by the manufacturer (Birkedal et al. 2012). However it contains 9% organic impurities. Therefore it is considered coated for the purpose of this case study and part of the respective group of nanoforms.

Link between structural (physicochemical) similarity and predicted property (C.3)

For NMs, the similarity cannot be based on molecular structure as for conventional chemicals. Moreover, in general very little is known about the mechanisms of toxic action of NMs. Therefore it can be challenging to develop the grouping hypothesis of the link between similar properties and the predicted property/toxicity, in this case the outcome of the *in vitro* comet assay.

The link between physicochemical properties of the analogues and the predicted property, i.e. genotoxicity, has been investigated with chemoinformatics methods. The hypothesis is that nano-TiO_2 in its uncoated form has the potential to damage DNA, but this can be masked by the presence of coating.

Consistency of effects in the data matrix (C.4), Reliability and adequacy of the source studies (C.5)

Data were collected following a reliability assessment based on the ANSES criteria.

The data for micronucleus, comet and chromosomal aberration assays were included in the data collection, because the current OECD test guidelines fort these tests are considered applicable to NMs and thus sufficiently reliable. For the *in vitro* comet assay however, an OECD test guideline is not available. This endpoint was considered nevertheless in the data collection because of larger availability of studies. Previous reviews on genotoxicity tests applied to NMs claimed that the comet and micronucleus assays are the most commonly used tests in the field (Golbamaki et al. 2015).

Generally, the assessment of quality, reliability and relevance to human health endpoints of measured toxicity data as well as their interpretation is difficult, partly due to the uncertainty in applying existing testing protocols to nanomaterials. The applicability of current OECD test guidelines to NMs is still under discussion and artefacts affecting the results of toxicity assessment of NMs have been reported (see Marchese Robinson et al. 2016).

Bias in selection of category members (C.6)

Because of the scarcity of the data available (as full datasets covering all properties considered) or insufficient quality, the set of analogues considered in this case study was limited. Only the nanoforms that were completely identified by means of fundamental parameters like hydrophobicity, zeta potential, dispersability were considered. This led to a dataset with only six nano-TiO_2 materials, differing in their primary particle size (from 7 to 117 nm), coating (two of them are declared coated by the manufacturer and the others are declared without a coating), crystal type (anatase and rutile) and hydrophobicity (materials functionalised to be hydrophobic or hydrophilic). Thus, the hypothesis is based on a small dataset and considers a single *in vitro* endpoint (comet assay). Other *in vitro* tests available for nano-TiO_2 do not confirm genotoxicity, so an overall conclusion on genotoxic potential cannot be made with certainty.

Compound the test organism is exposed to (6.1)

Nano-TiO2 without coating is considered to cause the adverse effect, the presence of coating is hypothesised to mask the potential DNA damage. The definition of the toxicant is based on evaluation of the physicochemical properties. Some uncertainty remains about the property driving the genotoxicity.

Common underlying mechanism, qualitative and quantitative aspects (6.2, 6.3)

The mechanisms of primary and secondary genotoxicity are not fully understood.

Generally, there is some uncertainty related to the mechanism: The majority of studies supported the hypothesis that the genotoxic effect of TiO_2 is masked by the presence of coating. However, the way in which the coating can prevent DNA damage is not entirely clear. The mechanism of genotoxicity of TiO_2 is not well defined and still discussed in the

literature; it is also possible that several effects take place at the same time. There is general uncertainty about the mechanism of action for indirect primary genotoxicity via ROS (Golbamaki et al. 2015), a clear correlation between the level of ROS production and DNA damage was not supported in several studies. Therefore measured information on ROS formation (bioactivity) would be useful for supporting the hypothesis in the case study.

Exposure to other compounds (6.4)

Exposure to impurities might influence the observed genotoxic effects. For example the presence as impurities or in the NM composition of reactive transition metals may also contribute to oxidative DNA damage induction.

Another factor to take into consideration is the presence of proteins in the medium. If the NMs are surrounded by proteins, they are more dispersed and also less toxic since the "reactive" part is "hidden" behind the protein corona.

Occurrence of other effects (6.5)

There might be more than one mechanism responsible for genotoxicity of TiO_2; and possibly a combination of several factors is responsible for masking the DNA damage. Generally, the presence of some form of coating may either prevent the aggregation of NMs, or the physical contact with DNA and/or other cell components, both with the result of reducing the positive results in the comet assay.

There is also an indication for the fact that the degree of agglomeration of nano-TiO_2 has an influence on the DNA damage.

A direct interaction mechanism of genotoxicity or an indirect primary genotoxicity are considered (Magdolenova et al. 2014). Furthermore, the conduction band of TiO_2 falls in the range of biological redox potentials (Burello and Worth 2011), meaning that TiO_2 with or without the presence of UV light can generate reactive species that may react with cell constituents such as DNA. However, the band gap is not the only predictor of reactivity, but also for example the geometry of the crystal or the nature of surface defects.

Summary of uncertainties in the case studies

The major uncertainties encountered in the nanomaterial case studies were related to:

- Complexity of nanostructures: similarity, category boundaries and members
- Identification of nanomaterials
- Quality and inconsistency/reproducibility of study data; missing protocols or uncertainty in the applicability of protocols to nanomaterials, both for material characterisation and toxicity assays
- Physicochemical properties driving the toxicity
- Limited datasets
- Read-across of negative effects
- Possible combination of physicochemical properties affecting toxicity

- Relevant test systems (*in vitro* mechanistic studies).

7.2. Nanospecificities to be considered in the RAAF

The following nanospecific issues have been identified related to a read-across for nanomaterials:

- Similarity based on chemical structure to be replaced with appropriate and relevant other parameter(s), i.e. consideration of physical form and key physicochemical properties

- Definition of the "identical or different" compounds in the RAAF should be adapted to nanoforms and consider factors such as coating, size, and length

- Definition of the compound the organism is exposed to and leading to an adverse effect:
  In the case of nanomaterials it is not only a distinction between parent compounds and (bio)transformation products such as metabolites, but for example needs to consider either the nanomaterial as such, or impurities/coating, released metals.

- Are there specific nanotoxicity mechanisms or kinetics of exposure which have to be taken into consideration? High variability of the measurements and possible nanospecific artefacts in assays are issues specific to NMs to be considered in the assessment.

The question of similarity is a specific fundamental issue. The RAAF, as related to the REACH legislation, is anchored on similarity of chemical structures.

For nanomaterials, the first step of defining a group of similar substances, i.e. nanoforms, has to be approached differently and similarity has to be defined based on the essential and relevant properties of the NMs. There is not a generally accepted way of defining similarity between NMs. The different frameworks that are available seem to agree on the fact that the similarity of NMs depends on the endpoints or uses considered. It has been reported that the importance of physicochemical data for NM are dependent on the type of NMs (Sellers et al. 2015).
Composition can be crucial for some applications and it could allow grouping (e.g. carbonanotubes), however in other contexts coatings or biodurability (dissolution rates) can be more important as for example for NMs that leach ions (e.g. Ag, ZnO). In addition, there is lack of publicly available data for NMs what makes it difficult to determine if two NMs have similar properties. But most importantly there is a lack of Testing Guidelines that guarantee that the same kind of data (e.g. size distribution) are generated in the same conditions and are comparable. Currently, some data for NMs can be found in publications but closer inspection shows that the same type of data was usually measured with different protocols (solvents, sonication times, sonication types, etc.) and by different research groups. These variations in protocol raise the question of whether data generated in different conditions are comparable. All this results into smaller datasets and data gaps, taking also into account that there are very few *in silico* tools that can be used to calculate properties for specific NMs. Thus, for nanomaterials the RAAF Assessment Element considerations based on chemical structure would need to be adapted appropriately.

Furthermore, the variability of the measurement data of the NM characterisation hampers linking the characteristics to the observed effects. In the literature, an appropriate definition

of the requirements for characterisation is repeatedly called for (see e.g. Nel et al. 2015; Fadeel et al. 2015; Marchese Robinson et al. 2016).

In general, however, the case study has shown that the RAAF framework is applicable to nanomaterials as well as to conventional chemicals.

8. CONSIDERATIONS FOR NANOMATERIAL READ-ACROSS

The case study has been documented by providing mechanistic interpretation of the available data, where possible, and according to the state of the art in the field. However, it has also to be acknowledged that the dataset for application of read-across to NMs was quite limited due to the data available and the set of analogues was also limited.

Applicability of the ECHA workflow

- The workflow proposed in the ECHA guidance for grouping and read-across of NMs (ECHA 2017a) was taken as basis for the nano-TiO_2 case study. It generally allowed to present all the information collected and used to build and test the grouping hypothesis. It was adapted slightly for the case study.

Relevance of computational methods in grouping for read-across

- It has been shown how unsupervised techniques like hierarchical clustering and PCA can support the grouping hypothesis by identifying the differences between nanoforms and by supporting the weight of evidence in the read-across hypothesis.

- Supervised and unsupervised techniques were applied in the case study presented. Unsupervised techniques are more useful early in the process for hypothesis formulation, to identify which properties are most relevant for grouping and the similarity between analogues on the basis of these properties. Supervised techniques are more useful later in the process, for model building when there is already evidence of property-activity relationships. In this case study, a supervised variable selection algorithm was used to determine the most valuable properties in predicting genotoxicity, to support the grouping hypothesis. Irrespective of the techniques used, they should be considered as feeding information into an overall weight of evidence, rather than being conclusive.

Uncertainty in grouping NMs for read-across

- Uncertainties associated with the read-across argumentation have been identified; they are related to the identification of the (non-)nanoforms, experimental variability associated with the physicochemical and toxicological information and due to the lack of measurement protocols for NMs, and, finally, to the lack of knowledge on the mechanisms of genotoxic action of NMs.

- A RAAF scenario was applied for the uncertainty analysis. A table of uncertainties was compiled relating to the ECHA RAAF Assessment Elements.

- The RAAF is generally applicable to NMs. A key aspect that would need to be extended for its application to NMs is the concept of similarity, as in the RAAF (following REACH) it is based on structural similarity. For NMs other principles for similarity should be included.

ACKNOWLEDGEMENT

The case study was part of a report produced by the European Commission's Joint Research Centre (JRC) for DG Internal Market, Industry, Entrepreneurship and SMEs (GROW) under the terms of Administrative Arrangement n. 33269.

9. REFERENCES

ANSES, 2016. *CLH report Proposal for Harmonised Classification and Labelling Substance Name : Titanium dioxide*, Available at: http://echa.europa.eu/harmonised-classification-and-labelling-previous-consultations/-/substance-rev/13832/term.

Azqueta, A. & Dusinska, M., 2015. The use of the comet assay for the evaluation of the genotoxicity of nanomaterials. *Frontiers in Genetics*, 6, p.239.

Barillet, S. et al., 2010. Toxicological consequences of TiO2, SiC nanoparticles and multi-walled carbon nanotubes exposure in several mammalian cell types: an *in vitro* study. *Journal of Nanoparticle Research*, 12(1), pp.61–73.

Birkedal, R. et al., 2012. *Nanogenotox deliverable 4.3: Crystallite size, mineralogical and chemical purity of NANOGENOTOX nanomaterials*, Copenhagen. Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

Burello, E. & Worth, A.P., 2011. A theoretical framework for predicting the oxidative stress potential of oxide nanoparticles. *Nanotoxicology*, 5(2), pp.228–35.

ECHA, 2017a. Guidance on information requirements and chemical safety assessment *Appendix R.6-1 for nanomaterials applicable to the Guidance on QSARs and Grouping of Chemicals*, Available at: https://echa.europa.eu/documents/10162/23036412/appendix_r6_nanomaterials_en.pdf/71ad76f0-ab4c-fb04-acba-074cf045eaaa.

ECHA, 2017b. *Read-Across Assessment Framework (RAAF)*, Helsinki. Available at: https://echa.europa.eu/documents/10162/13628/raaf_en.pdf.

ECHA, 2017c. *Appendix 4 : Recommendations for nanomaterials applicable to the Guidance on Registration Draft (Public) Version 1.0*.

ECHA, 2017d. *Guidance for identification and naming of substances under REACH and CLP*. Available at: https://echa.europa.eu/documents/10162/23036412/substance_id_en.pdf/ee696bad-49f6-4fec-b8b7-2c3706113c7d.

Fadeel, B. et al., 2015. Keeping it real: The importance of material characterization in nanotoxicology. *Biochemical and Biophysical Research Communications*, 468(3), pp.498–503.

Falck, G. et al., 2009. Genotoxic effects of nanosized and fine TiO2. *Human & Experimental Toxicology*, 28(6–7), pp.339–352.

Golbamaki, N. et al., 2015. Genotoxicity of metal oxide nanomaterials: review of recent data and discussion of possible mechanisms. *Nanoscale*, 7(6), pp.2154–2198..

Guichard, Y. et al., 2012. Cytotoxicity and genotoxicity of nanosized and microsized titanium dioxide and iron oxide particles in Syrian hamster embryo cells. *The Annals of occupational hygiene*, 56(5), pp.631–44.

Husson, F. et al., 2011, Exploratory multivariate analysis by example using R, CRC Press.

Jensen, K.A., Kembouche, Y. & Nielsen, S.H., 2013. *Nanogenotox deliverable 4.7: Hydrochemical reactivity, solubility, and biodurability of NANOGENOTOX nanomaterials*, Copenhagen. Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

Jugan, M.-L. et al., 2012. Titanium dioxide nanoparticles exhibit genotoxicity and impair DNA repair activity in A549 cells. *Nanotoxicology*, 6(5), pp.501–513.

Liaw, A., Wiener, M.,2002, Classification and Regression by randomForest. *R News*, 2(3), pp. 18–22.

Karlsson, H.L. et al., 2015. Can the comet assay be used reliably to detect nanoparticle-induced genotoxicity? *Environmental and Molecular Mutagenesis*, 56(2), pp.82–96.

Li, K. et al., 2013. Nanoparticles Inhibit DNA Replication by Binding to DNA: Modeling and Experimental Validation. *ACS Nano*, 7(11), pp.9664–9674.

Luttrell, T. et al., 2014. Why is anatase a better photocatalyst than rutile? - Model studies on epitaxial TiO2 films. *Scientific Reports*, 4, p.4043.

Magdolenova, Z. et al., 2012. Impact of agglomeration and different dispersions of titanium dioxide nanoparticles on the human related *in vitro* cytotoxicity and genotoxicity. *Journal of Environmental Monitoring*, 14(2), p.455.

Magdolenova, Z. et al., 2014. Mechanisms of genotoxicity. A review of *in vitro* and *in vivo* studies with engineered nanoparticles. *Nanotoxicology*, 8(3), pp.233–278.

Mano, S.S. et al., 2012. Effect of Polyethylene Glycol Modification of TiO 2 Nanoparticles on Cytotoxicity and Gene Expressions in Human Cell Lines. *International Journal of Molecular Sciences*, 13, pp.3703–3717.

Marchese Robinson, R.L. et al., 2016. How should the completeness and quality of curated nanomaterial data be evaluated? *Nanoscale,* 8, pp.9919-9943.

Nanogenotox, 2012. Towards a method for detecting the potential genotoxicity of nanomaterials D4.2: Transmission electron microscopic characterization of NANOGENOTOX nanomaterials, Final report Key intrinsic physicochemical characteristics of NANOGENOTOX nanomaterials WP. Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

Nanogenotox WP6, 2013. *Characterisation of manufactured nanomaterials for their clastogenic/ aneugenic effects or DNA damage potentials and correlation analysis*, Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

NanoGenoTox Joint Action, 2013. NANOGENOTOX Final report. Facilitating the safety evaluation of manufactured nanomaterials by characterising their potential genotoxic hazard. Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

Nel, A.E. et al., 2015. Where Are We Heading in Nanotechnology Environmental Health and Safety and Materials Characterization? *ACS Nano*, 9(6), pp.5627–5630.

Norppa, H. et al., 2013. *Nanogenotox deliverable 5 : In vitro testing strategy for nanomaterials including database Final report*, Helsinki. Available at: http://www.nanogenotox.eu/index.php?option=com_content&view=article&id=136&Itemid=158.

OECD, 2015. Titanium dioxide (NM100-NM105). Dossier Part1 to Part 5, Paris, Working Party on Manufactured Nanomaterials. Available at: http://www.oecd.org/chemicalsafety/nanosafety/titanium-dioxide-nm100-nm105-manufactured-nanomaterial.htm.

R CoreTeam, 2016. R: A language and environment for statistical computing.

Rasmussen, K. et al., 2016. Review of achievements of the OECD Working Party on Manufactured Nanomaterials' Testing and Assessment Programme. From exploratory testing to test guidelines. *Regulatory toxicology and pharmacology*, 74, pp.147–160.

Rasmussen, K. et al., 2014. *Titanium Dioxide, NM-100, NM-101, NM-102, NM-103, NM-104, NM-105: Characterisation and Physico- Chemical Properties*, Luxembourg.

Rice, Z. et al., 2009. Terminal phosphate group influence on DNA - TiO_2 nanoparticle interactions. *MRS Proceedings*, 1236, pp.1236-SS05-15..

RIVM, JRC & ECHA, 2016. *Usage of (eco)toxicological data for bridging data gaps between and grouping of nanoforms of the same substance. Elements to consider*, Available at: https://echa.europa.eu/documents/10162/13630/eco_toxicological_for_bridging_grouping_nanoforms_en.pdf.

Romoser, A.A., 2012. *Cytotoxicological response to engineered nanomaterials: A pathway-driven process PhD Dissertation. Available at: http://oaktrust.library.tamu.edu/bitstream/handle/1969.1/ETD-TAMU-2012-05-10877/ROMOSER-DISSERTATION.pdf?sequence=2*.

SCCS, 2014. *Opinion on Titanium Dioxide (nano form), COLIPA n° S75, revision of 22 April 2014*, Brussels: Scientific Committee on Consumer Safety, SCCS/1516/13. DOI:10.2772/70108.

Schröder, K. et al., 2014. Carcinogenicity and Mutagenicity of Nanoparticles – Assessment of Current Knowledge as Basis for Regulation. Umweltbundesamt, Texte 50/2014. Available at: http://www.umweltbundesamt.de/publikationen/carcinogenicity-mutagenicity-of-nanoparticles.

Schultz, T.W. et al., 2015. A Strategy for Structuring and Reporting a Read-Across Prediction of Toxicity, *Regulatory toxicology and pharmacology, 72, pp.586–601*.

Sellers, K. et al., 2015. Grouping nanomaterials A strategy towards grouping and read-across, RIVM Report 2015-0061. Available at: http://www.rivm.nl/en/Documents_and_publications/Scientific/Reports/2015/juni/Grouping_nanomaterials_A_strategy_towards_grouping_and_read_across.

Stone, U.K.V. et al., 2014. ITS-NANO - Prioritising nanosafety research to develop a stakeholder driven intelligent testing strategy, *Particle and Fibre Toxicology*, 11:9.

Suzuki, R., Shimodaira, H., 2006. Pvclust: an R package for assessing the uncertainty in hierarchical clustering, *Bioinformatics*, 22, pp. 1540–1542.

Trouiller, B. et al., 2009. Titanium dioxide nanoparticles induce DNA damage and genetic instability *in vivo* in mice. *Cancer research*, 69(22), pp.8784–8789.

Worth, A. et al., 2017. Evaluation of the availability and applicability of computational approaches in the safety assessment of nanomaterials. Final report of the Nanocomput project. EUR28617EN. DOI: 10.2760/248139.

ANNEX

Table A1. Full dataset used for the nano-TiO_2 read-across (for the clustering, principal component analysis, random forest).

Name	*NM-100*	*NM-101*	*NM-102*	*NM-103*	*NM-104*	*NM-105*
In vitro comet assay	1	0	1	0	0	1
Total non-TiO_2 content including coating and impurities (% w/w)	1.5	9	5	11	11	0.11
Impurity (% w/w Fe)	0.49	0	0.07	0.06	0	0.06
Impurity (% w/w Si)	0.28	0.29	0.08	0.68	0.018	0.07
Impurity (% w/w K)	0.25	0	0.001	0.001	0.001	0
Impurity (% w/w P)	0.21	0.27	0.001	0	0	0
Impurity coating (% w/w Al)	0.09	0.09	0.05	3.4	3.2	0.04
Impurity (% w/wCr)	0.03	0	0	0	0	0
Impurity (% w/w Zr)	0.005	0.01	0.005	0.001	0.001	0
Impurity (% w/w Ca)	0.001	0	0.005	0.005	0.01	0
Impurity (% w/w Na)	0.001	0.1	0.001	0.01	0	0.001
Impurity (% w/w S)	0	0.22	0.001	0.01	0.01	0.26
Impurity (% w/w Mg)	0	0	0	0.001	0.001	0
Organic matter (% w/w)	0	8	0	2	2	0
Crystal type (Anatase)	1	1	1	0	0	0.84
Crystal type (Rutile)	0	0	0	1	1	0.16
Crystal type (Cubic)	0	0	0	0	0	0
Crystallite size (mean)	117.81	7.69	23.93	24.32	24.71	22.44
Surface coating (declared) (%)	0	0	0	8	8	0
Specific surface area (m^2/g)	9.23	316.07	77.86	53.98	54.33	47.00
Shape (elongated='1,' spherical=0)	0	0	0	1	0	1
Aspect ratio	1.53	1.53	1.53	1.70	1.53	1.36
Primary particle diameter (mean)	93.45	5.25	22.00	24.00	24.50	20.13
Particle size distribution (Z-average) (nm)	210	278	439.8	135.11	144.47	176.78
Particle size distribution-SD (Z-average) (nm)	10	0	36.66	25.27	35.21	38.99
Particle size distribution (Intensity distribution peak (nm))	NA	NA	685.55	146.62	193.94	180.75
Particle size distribution-SD (Intensity distribution peak (nm))	NA	NA	30.8	21.46	48.36	17.98
Particle size distribution (FWHM main peak)	NA	NA	444.5	82.36	100.52	75.47
Particle size distribution-SD (FWHM main peak)	NA	NA	94.9	12.74	31.37	10.20
Particle size distribution (after 24h) (nm)	NA	NA	969	198	NA	214
Particle size distribution-SD (after 24h) (nm)	NA	NA	7.65	NA	NA	NA
Particle size distribution in MQ Water, untreated, Mode #1 (nm)	391.2	1609	115	973.2	727.8	1102
Particle size distribution in MQ Water, untreated, Mode #2 (nm)	4862	1609	5170	973.2	727.8	204.7
Particle size distribution in PBS, untreated, Mode #1 (nm)	1440	1188	1528	1977	1817	4526
Particle size distribution in PBS, untreated, Mode #2 (nm)	5236	5148	5330	1977	5194	1150
Particle size distribution in DMEM + Lglutamate, untreated, Mode #1 (nm)	995.5	1438	2745	2255	3059	1881
Particle size distribution in DMEM + Lglutamate, untreated, Mode #2 (nm)	995.5	5560	2745	2255	3059	5372
Particle size distribution in DMEM + 1% FBS, untreated, Mode #1 (nm)	736	1201	1415	1040	1156	2454
Particle size distribution in DMEM + 1% FBS, untreated, Mode #2 (nm)	736	5232	1415	4593	5211	626.5
Particle size distribution in DMEM + 5% FBS, untreated, Mode #1 (nm)	845.4	1278	1414	991.1	719.3	1709
Particle size distribution in DMEM + 5% FBS, untreated, Mode #2 (nm)	845.4	1278	1414	991.1	5375	1709
Particle size distribution in DMEM + 10%FBS, untreated, Mode #1 (nm)	639.1	1406	1521	1156	711.2	1030
Particle size distribution in DMEM + 10%FBS, untreated, Mode #2 (nm)	4793	1406	1521	1156	711.2	4731
Particle size distribution in MQ water, untreated (Z-average) (nm)	343	1746	1062	671.6	367.8	720
Particle size distribution in PBS , untreated (Z-average) (nm)	2289	1229	1579	1397	1600	3342
Particle size distribution in DMEM + Lglutamine untreated (Z-average) (nm)	2129	1954	2427	1665	2869	2868
Particle size distribution in DMEM + 1% FBS,untreated (Z-average) (nm)	606.8	1166	1295	828.8	1111	1599
Particle size distribution in) DMEM + 5% FBS, untreated (Z-average) (nm)	621.3	1039	1234	653.2	657.5	1116
Particle size distribution in DMEM + 10% FBS, untreated (Z-average) (nm)	582.4	1127	1227	683.3	617.8	937.3
Particle size distribution in MQ Water, 1min sonication, Mode #1 (nm)	259.3	719.5	703	2649	207.7	352.6
Particle size distribution in MQ Water, 1min sonication, Mode #2 (nm)	259.3	719.5	703	2649	207.7	352.6
Particle size distribution in PBS, 1min sonication, Mode #1 (nm), Mode #1	2116	2254	2525	1629	4031	1682

Particle size distribution in PBS,1min sonication, Mode #2 (nm)	2116	2254	2525	4619	465.2	5108
Particle size distribution in DMEM + Lglutamate, 1min sonication, Mode #1 (nm)	2973	2854	3488	4043	1701	4673
Particle size distribution in DMEM + Lglutamate, 1min sonication, Mode #2 (nm)	2973	2854	3488	4043	5560	1995
Particle size distribution in DMEM + 1% FBS , 1min sonication, Mode #1 (nm)	405.3	678.5	837.5	275.6	333.6	306.8
Particle size distribution in DMEM + 1% FBS, 1min sonication, Mode #2 (nm)	405.3	678.5	189.8	4344	4670	4755
Particle size distribution in DMEM + 5% FBS , 1min sonication, Mode #1 (nm)	408.8	755.5	901.8	432.4	278.2	336.9
Particle size distribution in DMEM + 5% , 1min sonication, Mode #2 (nm)	408.8	755.5	115.6	4881	278.2	4755
Particle size distribution in DMEM + 10%FBS, 1min sonication, Mode #1 (nm)	345.8	823.6	1077	370.9	334.4	349.8
Particle size distribution in DMEM + 10%FBS, 1min sonication, Mode #2 (nm)	345.8	823.6	1077	370.9	334.4	349.8
Particle size distribution in MQ water, 1min sonication (Z-average) (nm)	201.3	500.9	505.7	1977	194.3	227.5
Particle size distribution in PBS, 1min sonication (Z-average) (nm)	1624	1827	2079	2275	3197	3585
Particle size distribution in DMEM + , 1min sonication (Z-average) (nm)	2514	2350	2701	3551	3306	3507
Particle size distribution in DMEM + 1% FBS - 1min sonication) (Z-average) (nm)	310.4	521.2	590	263.5	278.5	265.3
Particle size distribution in DMEM + 5% FBS - 1min sonication) (Z-average) (nm)	315.2	569.2	617.3	345.8	225.8	286.3
Particle size distribution in DMEM + 10% FBS - 1min sonication) (Z-average) (nm)	283.9	623.4	732.2	286.9	267.8	281.2
Particle size distribution in MQ Water, 20min sonication, Mode #1 (nm)	378.8	1111	1103	765.3	344.5	902
Particle size distribution in MQ Water , 20min sonication, Mode #2 (nm)	378.8	4077	193.6	5041	4638	243.9
Particle size distribution in PBS , 20min sonication, Mode #1 (nm)	1042	1265	1789	1449	2779	4437
Particle size distribution in PBS Mode #2, 20min sonication, Mode #2 (nm)	5236	4976	4988	5037	2779	4437
Particle size distribution in DMEM + Lglutamate, 20min sonication, Mode #1 (nm)	1059	1974	2001	2916	3207	1956
Particle size distribution in DMEM + Lglutamate, 20min sonication, Mode #2 (nm)	1059	4881	5517	2916	3207	5290
Particle size distribution in DMEM + 1% FBS, 20min sonication, Mode #1 (nm)	631.9	1368	1063	684.1	975.4	969.6
Particle size distribution in DMEM + 1% FBS, 20min sonication, Mode #2 (nm)	5059	420.5	1063	4946	286.9	226.6
Particle size distribution in DMEM + 5% FBS, 20min sonication, Mode #1 (nm)	522.7	1073	1487	1079	925	848.3
Particle size distribution in DMEM + 5% FBS, 20min sonication, Mode #2 (nm)	5017	5046	1487	1079	925	4864
Particle size distribution in DMEM + 10%FBS, 20min sonication, Mode #1 (nm)	565.7	1255	1228	1155	605.2	1110
Particle size distribution in DMEM + 10%FBS, 20min sonication, Mode #2 (nm)	565.7	5222	1228	262.3	4991	1110
Particle size distribution in MQ water - 20min sonication)	307.6	1130	794.1	596.9	290.8	474.4
Particle size distribution in PBS,- 20min sonication (Z-average) (nm)	1217	1276	1809	1350	2284	4514
Particle size distribution in DMEM + Lglutamine, 20min sonication (Z-average) (nm)	1754	1992	1997	2268	2636	2938
Particle size distribution in DMEM + 1% FBS, 20min sonication (Z-average) (nm)	540.2	668.7	975.4	526.8	520.4	743.9
Particle size distribution in DMEM + 5% FBS, 20min sonication (Z-average) (nm)	450.4	1065	1197	656.9	696.2	921.8
Particle size distribution in DMEM + 10% FBS, 20min sonication (Z-average) (nm)	473.1	957.9	874.8	570.3	480.8	619.8
Zeta Potential in MQ Water, 1min sonication (mV)	-24.5	-27.2	-27.1	39.1	-23.4	-23.8
Zeta Potential in PBS, 1min sonication (mV)	-26.7	-19.7	-25.1	-20.8	-16.9	-20.5
Zeta Potential in DMEM + Lglutamate, 1min sonication, Mode #1 (mV)	20.5	22.3	-3.14	-8.44	-7.29	-2.55
Zeta Potential in DMEM + Lglutamate, 1min sonication, Mode #2 (mV)	-26	-34.3	-3.14	-8.44	-7.29	-2.55
Zeta Potential in DMEM + Lglutamate, 1min sonication, Mode #3 (mV)	95.4	-92	-3.14	-8.44	-7.29	-2.55

Zeta Potential in DMEM + 1% FBS, 1min sonication, Mode #1 (mV)	-9.14	-11.8	-13.6	-9.98	-8.88	-9.37
Zeta Potential in DMEM + 1% FBS, 1min sonication, Mode #2 (mV)	140	-11.8	-13.6	-9.98	-8.88	-9.37
Zeta Potential in DMEM + 5% FBS, 1min sonication, Mode #1 (mV)	107	-15	-13.4	-12	15.1	9.43
Zeta Potential in DMEM + 5% FBS, 1min sonication, Mode #2 (mV)	35.4	-15	-13.4	-12	-43.7	-47.9
Zeta Potential in DMEM + 5% FBS, 1min sonication, Mode #3 (mV)	-21.2	-15	-13.4	-12	-43.7	-47.9
Zeta Potential in DMEM + 10%FBS, 1min sonication, Mode #1 (mV)	78.4	0.13	-10.5	-12.4	-9.38	-9.92
Zeta Potential in DMEM + 10%FBS, 1min sonication, Mode #2 (mV)	12.2	0.13	-10.5	-12.4	129	-9.92
Zeta Potential in DMEM + 10%FBS, 1min sonication, Mode #3 (mV)	-27.1	0.13	-10.5	-12.4	129	-9.92
Zeta Potential in MQ Water - 20min sonication (mV)	-40.6	-27.5	30.3	39.1	24.6	-32.6
Zeta Potential in PBS - 20min sonication (mV)	-20.2	-21.7	-18.5	-20.9	-20.3	-33.2
Zeta Potential in DMEM + Lglutamate, 20min sonication, Mode #1 (mV)	-1.55	3.6	-3.46	-8.76	-9.98	-8.55
Zeta Potential in DMEM + Lglutamate, 20min sonication, Mode #2 (mV)	-1.55	-42.5	-3.46	-8.76	-9.98	-8.55
Zeta Potential in DMEM + 1% FBS, 20min sonication (mV)	-11.4	-12	-12.4	-10	-10.2	-7.76
Zeta Potential in DMEM + 5% FBS, 20min sonication (mV)	-10.4	-11.3	-9.47	-13.7	-9.38	-11.9
Zeta Potential in DMEM + 10%FBS, 20min sonication, Mode #1 (mV)	-11.3	-11.5	-10.4	-11.8	-10.5	-5.43
Zeta Potential in DMEM + 10%FBS, 20min sonication, Mode #2 (mV)	-11.3	-11.5	-10.4	-11.8	-10.5	-74
Polydispersity Index in MQ water - untreated	0.205	0.264	0.187	0.287	0.376	0.376
Polydispersity Index in PBS - untreated	0.219	0.239	0.769	0.255	0.232	0.232
Polydispersity Index in DMEM + Lglutamine untreated	0.332	0.359	0.181	0.256	0.247	0.247
Polydispersity Index in DMEM + 1% FBS - untreated	0.207	0.201	0.081	0.269	0.208	0.208
Polydispersity Index in DMEM + 5% FBS - untreated	0.194	0.232	0.139	0.293	0.22	0.22
Polydispersity Index in DMEM + 10% FBS - untreated	0.176	0.194	0.182	0.369	0.201	0.201
Polydispersity Index in MQ water - 1min sonication	0.205	0.274	0.248	0.393	0.236	0.211
Polydispersity Index in PBS - 1min sonication	0.219	0.283	0.188	0.442	0.334	0.443
Polydispersity Index in DMEM + Lglutamine - 1min sonication	0.332	0.217	0.268	0.279	0.434	0.395
Polydispersity Index in DMEM + 1% FBS - 1min sonication	0.207	0.232	0.243	0.243	0.194	0.177
Polydispersity Index in DMEM + 5% FBS - 1min sonication	0.194	0.232	0.27	0.25	0.161	0.207
Polydispersity Index in DMEM + 10% FBS - 1min sonication	0.176	0.24	0.27	0.196	0.178	0.196
Polydispersity Index in MQ water - 20min sonication	0.199	0.351	0.254	0.393	0.306	0.443
Polydispersity Index in PBS - 20min sonication	0.317	0.238	0.231	0.25	0.227	0.274
Polydispersity Index in DMEM + Lglutamine - 20min sonication	0.515	0.247	0.227	0.264	0.209	0.341
Polydispersibility Index in DMEM + 1% FBS - 20min sonication	0.195	0.282	0.054	0.317	0.282	0.48
Polydispersibility Index in DMEM + 5% FBS - 20min sonication	0.223	0.302	0.179	0.367	0.221	0.456
Polydispersibility Index in DMEM + 10% FBS - 20min sonication	0.204	0.234	0.235	0.417	0.239	0.391
Polydispersity Index (PdI)	NA	0.323	0.427	0.292	0.227	0.245
IsoelectricPoint (Mean)	NA	5.5	6	8.3	8.5	6.8
IsoelectricPoint (Min)	NA	5.3	6	8.2	8.2	6.6
IsoelectricPoint (Max)	NA	5.7	6	8.5	8.8	6.9
Density (g/mL)	3.84	3.99	3.84	4.015	4.09	4.052
Mean of total pore volume (mL/g)	0.0324	0.319	0.2996	0.2616	0.1935	0.1937
Micro surface area (m^2/g)	0	13.625	1.108	0	0	0
Micropore volume (mL/g)	0	0.00179	0.00034	0	0	0
Specific surface area (mean)	9.23	242.785	77.864	53.983	54.331	47
Dustiness-Respirable(mg/kg)	1500	5600	9200	19000	6400	11000
Biodurability 24h 0.05% BSA (Ti content) (g/l)	5.2	0	0	0	0	0
Biodurability 24h Gambles solution (Ti content) (g/l)	0	0	3388	0	0	0
Biodurability 24h Caco2 (Ti content) (g/l)	796	3414	1741	222	3386	2724
Biodurability 24h 0.05% BSA (Al content) (g/l)	0	175	0	198	137	0
Biodurability 24h Gambles solution (Al content) (g/l)	0	177	0	868	413	0
Biodurability 24h Caco2 (Al content) (g/l)	24	252	0	182	413	0
Biodurability 24h 0.05% BSA (Si content) (mg/l)	0	0	0	0.9	0	0
Biodurability 24h Gambles solution (Si content) (g/l)	0	0	0	2.0	0	0
Biodurability 24h Caco2 (Si content) (g/l)	0	0	0	1.7	0	0
Redox Caco2 medium [a]	1	-1	-1	1	-1	-1

Redox Gamble's solution [a]	1	0	-1	1	-1	-1
Redox BSA [a]	0	0	0	0	0	0

a values obtained from Nanogenotox 4.7 determined by measuring the content of O2. Oxidising properties (1), neutral (0), reducing (-1)

Table A2. References for the *in vitro* comet assay results for nano-TiO_2 NMs taken into account in

Table A2. References for the *in vitro* comet assay results for nano-TiO_2 NMs taken into account in the case study.

Type the subtitle here. If you do not need a subtitle, please delete this line.

	Assay outcome (N of positives/ total N studies)	References
NM 100	2/2	Norppa et al 2013
NM 101	2/6	Gerloff et al. 2012, Kermanizadeh et al 2013, 2012
NM 102	5/8	Kansara et al 2015, Vales et al 2015
NM 103	0/6	Norppa et al 2013
NM 104	0/6	Norppa et al 2013
NM 105	10/14	Armand et al 2016, Barillet et al 2009, Gerloff et al 2012, Guichard et al 2012, Jugan et al 2012, Patel et al 2014, Stoccoro et al 2016

Armand, L., Tarantini, A., Beal, D., Biola-Clier, M., Bobyk, L., Sorieul, S., Pernet-Gallay, K., Marie-Desvergne, C., Lynch, I., Herlin-Boime, N., Carriere, M., 2016. Long-term exposure of A549 cells to titanium dioxide nanoparticles induces DNA damage and sensitizes cells towards genotoxic agents, *Nanotoxicology*, 10, pp.913–923.

Barillet, S., Jugan, M.L., Simon-Deckers, A., Leconte, Y., Herlin-Boime, N., Mayne-L'Hermite, M., Reynaud, C., Carrière, M., 2009. SiC nanoparticles cyto- and genotoxicity to Hep-G2 cells, in: *Nanosafe 2008: International Conference on Safe Production and Use of Nanomaterials, Journal of Physics: Conference Series*, 170, 012016. https://doi.org/10.1088/1742-6596/170/1/012016

Gerloff, K., Fenoglio, I., Carella, E., Kolling, J., Albrecht, C., Boots, A.W., Förster, I., Schins, R.P.F., 2012. Distinctive Toxicity of TiO2 Rutile/Anatase Mixed Phase Nanoparticles on Caco-2 Cells, *Chemical Research in Toxicology*, 25, pp.646–655.

Guichard, Y., Schmit, J., Darne, C., Gaté, L., Goutet, M., Rousset, D., Rastoix, O., Wrobel, R., Witschger, O., Martin, A., Fierro, V., Binet, S., 2012. Cytotoxicity and genotoxicity of nanosized and microsized titanium dioxide and iron oxide particles in Syrian hamster embryo cells, *Ann. Occup. Hyg.*, 56, pp.631–44.

Jugan, M.-L., Barillet, S., Simon-Deckers, A., Herlin-Boime, N., Sauvaigo, S., Douki, T., Carriere, M., 2012. Titanium dioxide nanoparticles exhibit genotoxicity and impair DNA repair activity in A549 cells, *Nanotoxicology*, 6, pp.501–513.

Kansara, K., Patel, P., Shah, D., Shukla, R.K., Singh, S., Kumar, A., Dhawan, A., 2015. TiO_2 nanoparticles induce DNA double strand breaks and cell cycle arrest in human alveolar cells, *Environ. Mol. Mutagen.*, 56, pp.204–217.

Kermanizadeh, A., Pojana, G., Gaiser, B.K., Birkedal, R., Bilanicova, D., Wallin, H., Jensen, K.A., Sellergren, B., Hutchison, G.R., Marcomini, A., Stone, V., 2012. *In vitro* assessment of engineered nanomaterials using a hepatocyte cell line: Cytotoxicity, pro-inflammatory cytokines and function markers, *Nanotoxicology*, 7(3), pp.301–313.

Kermanizadeh, A., Vranic, S., Boland, S., Moreau, K., Baeza-Squiban, A., Gaiser, B.K., Andrzejczuk,

L.A., Stone, V., 2013. An *in vitro* assessment of panel of engineered nanomaterials using a human renal cell line: cytotoxicity, pro-inflammatory response, oxidative stress and genotoxicity, *BMC Nephrol.*, 14, 96.

Norppa, H., Siivola, K., Fessard, V., Tarantini, A., Apostolova, M., Jacobsen, N., Wallin, H., Goetz, M., Fieblinger, D., Stepnik, M., Simar, S., Quarre, S., Nesslany, F., De Jong, W.H., Marcos, R., Vales, G., Troisfontaines, P., Guichard, Y., Tavares, A., Louro, H., Silva, M., 2013. *Nanogenotox deliverable 5 : In vitro testing strategy for nanomaterials including database, Final report*, Helsinki.

Patel, R.M., Shete, P.B., Thorat, N.D., Otari, S.V., Barick, K.C., Prasad, A., Ningthoujam, R.S., Tiwale, B.M., Pawar, S.H., 2014. Superparamagnetic iron oxide/chitosan core/shells for hyperthermia application: Improved colloidal stability and biocompatibility, *J. Magn. Magn. Mater.*, 355, pp.22–30. https://doi.org/10.1016/j.jmmm.2013.11.033

Stoccoro, A., Di Bucchianico, S., Uboldi, C., Coppedè, F., Ponti, J., Placidi, C., Blosi, M., Ortelli, S., Costa, A.L., Migliore, L., 2016. A panel of *in vitro* tests to evaluate genotoxic and morphological neoplastic transformation potential on Balb/3T3 cells by pristine and remediated titania and zirconia nanoparticles, *Mutagenesis*, 31(5), pp.511–529.

Vales, G., Rubio, L., Marcos, R., 2015. Long-term exposures to low doses of titanium dioxide nanoparticles induce cell transformation, but not genotoxic damage in BEAS-2B cells, *Nanotoxicology*, 9, pp.568–578.