Testing for *in vitro* genetic toxicity in high dimensional nanomaterial dose-response experiments

Rahmasari Nur Azizah, Geert R. Verheyen, Sabine Van Miert and Ziv Shkedy

1. Introduction

Nanomaterials are materials with diameter in the range of 1 to 100 nm in at least one dimension (Jeevanandam et al., 2018). According to their origin, nanomaterials can be categorized into incidental, engineered and naturally produced nanomaterials. Incidental nanomaterials are formed incidentally in the industrial process, engineered nanomaterials are intentionally produced for their certain characteristics to be applied in desired products while naturally produced nanomaterials are nanomaterials that are present in human bodies, animals and plants (Jeevanandam et al., 2018). Natural nanomaterials have long been found on the earth in great abundance while incidental and engineered nanomaterials have been increasingly produced since the Industrial Revolution (Hochella et al., 2019). Nanomaterials, due to their valuable properties that differ from the same materials in bulk form, attract increasing interest and they are commonly developed and used (Yokel and MacPhail, 2011). Engineered nanomaterials occur in various industries such as food, medicine, sport, textile, cosmetics etc. They might appear in food packaging and skin care products, might be used as food additives, as nanofibres for wound-healing, or for drug delivery systems (Gubala et al., 2018). Nanomaterials are also used in agriculture industry as nanofertilizers, nanoparticles encapsulated pesticides, nanosensors for soil quality assessment, etc (Pramanik et al., 2020).

The many uses of engineered nanomaterials are mostly due to their useful mechanical, optical, magnetic and biological properties caused by their unique size-dependent properties that are not present in their bulkier form (Gubala et al., 2018). However, engineered nanomaterials with these enhanced properties might also be potentially dangerous (Van Miert, 2019). Studies found that general acute toxic effects such as reactive oxygen species generation and protein denaturation can be caused by nanomaterials exposure (Jeevanandam et al., 2018). Nanomaterials can have chronic toxic effects on possible organ enlargement and dysfunction as an effect of uptake by the reticuloendothelial system, nucleus, neuronal tissue and the generation of neoantigens. Nanoscale fibrous particles produced by the mining of asbestos can also cause a health risk if absorbed in the lungs (Jeevanandam et al., 2018). Nanomaterials may have negative effects on the environment as well, such as the presence of nanosilver in soil that may be uptaken by plants can lead to a reduction in the plants' growth, the effect of nano-ZnO on the growth of algae and the effect of nanomaterials on unicellular aquatic organisms and creatures in the aquatic life (Kabir et al., 2018). Thus, potential risks of nanomaterials need to be analyzed to ensure their safety for humans, plants, animals and the environment. The toxicity of the nanomaterials

In this paper we focus on the assessment of the risk of nanomaterial toxicity effects. This is done by developing a statistical methodology for risk assessment for toxicity effects in *in vitro* nanomaterials experiments and by developing (free and publicly available) software to conduct the analysis. A way to assess potential risks of nanomaterials is through performing dose-response analysis on the nanomaterial toxicity. Dose-response experiments involve a compound that is administered to a cell line, animal or human at several doses and a measured response. The relationship between the dose (concentration) and the observed response can then be evaluated through these experiments (Lin et al., 2012). For the analysis that is presented in this paper, dose is the amount of nanomaterial administered and response is the endpoint that indicates toxicity.

The database that is used for illustration in this study was constructed as a part of the H2020 NanoInformaTIX project and consists of *in vitro* dose-response data of 102 nanomaterials in several experimental settings. The aim of the analysis is to detect genetic toxicity *in vitro* data with DNA strand breaks as the endpoint of interest that indicates toxicity. In this paper several methods are presented to evaluate nanomaterials toxicity

with an underlying assumption that the true dose-response relationship between the dose (concentration) and the DNA strand breaks is monotone. As genetic toxicity for a large number of nanomaterials is tested, procedures to adjust for multiplicity are provided as well. In addition to methodology development, a new R package, `NMTox` is presented, in which the proposed methodology is implemented.
This paper is organized as follows. In Section 2 the *in vitro* experimental data, the experimental setting and the inclusion criteria for the analysis are presented. Methodology to detect *in vitro* nanomaterials toxicity is presented in Section 3. The proposed methodology is implemented and presented in Section 4 while the capacity of the `NMTox` R package is discussed in the section 5. Section 6 is devoted to a discussion of the analyses' results is given.

2. Data

2.1 Data Structure

2.1.1 Dose-response of NanoInformaTIX experiments

The data used as a case study in this paper originates from the NanoInformaTIX project (NanoInformaTIX, 2019) and it is stored in the eNanoMapper database (Jeliazkova, et al., 2015). The database contains data from several projects (eNanoMapper, n.d.) consisting of the results of toxicity and eco-toxicity studies and the results of experiments measuring physico-chemical properties of nanomaterials. The database contains the results of genetic toxicity endpoints obtained from *in vitro* studies and it consists of 30083 observations from 27 different projects that includes 102 nanomaterials with 57 types of endpoints. An example of the dose-response scatter plots of several endpoints for different nanomaterials is shown in Figure 1.

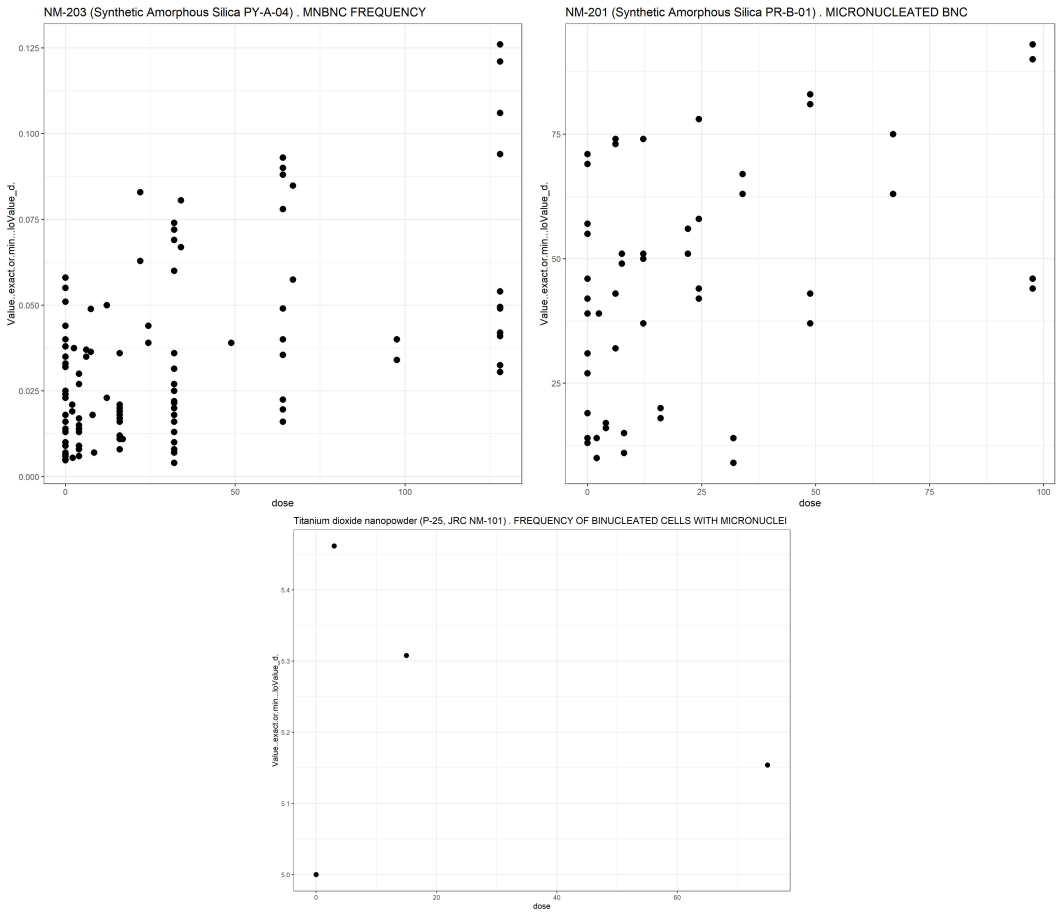


Figure 1: Scatter plot of several endpoints and nanomaterials in the dataset

Since genetic toxicity *in vitro* data contains data from several projects, it is possible to find slight differences in the way the data from different projects or laboratories were recorded. For example, among the endpoints listed in the dataset, there might be more than one endpoints that actually refer to the same type of endpoint. To identify this, user should refer the documentation of the related project.
For the analysis presented in this paper, the primary interest is to investigate genetic toxicity which causes DNA damage. A comet assay is used as a method to detect the DNA damage. A typical example of an *in vitro* dose-response experiment for cellulose 30000 nm is shown in Figure 2.

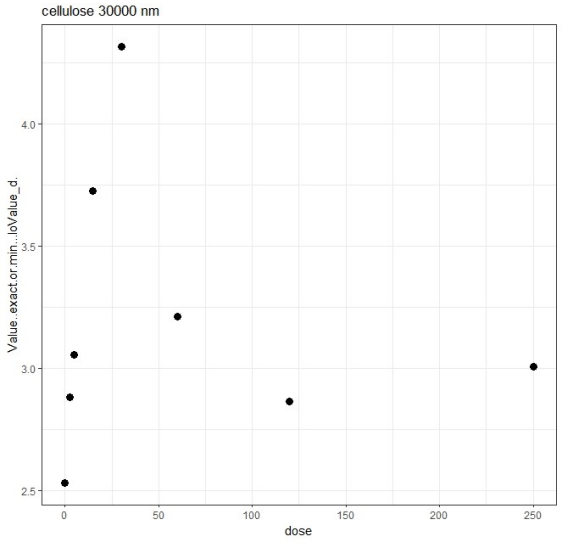


Figure 2: Dose-response plot of cellulose 30000 nm

The DNA damage is indicated by DNA strand breaks which is measured by the percentage of the DNA in the tail of a comet. In total there are 6816 observations with DNA strand breaks as an endpoint. These observations come from 75 nanomaterials and chemical substances. Since the interest is only on the nanomaterials, 20 chemical substances were omitted from the dataset, resulting in a final database with 55 nanomaterials. Suppose that a DNA damage for a specific nanomaterial m is denoted by $\mathbf{Y_m}$, then the observations for that particular nanomaterial can be illustrated as

$$\mathbf{Y_m} = \left(\; Y_{01m} \quad ... \quad Y_{0j1} \quad | \quad ... \quad | \quad ... \quad Y_{k_m n_i m} \; \right).$$

Here, the observation Y_{ijk}, denotes the DNA damage at the ith dose, $i = 0, \ldots, k_m$, for the J replicate, $j = 0, \ldots, n_i$ for the m^{th} for nanomaterial, $m = 1, \ldots, 55$.

2.1.2 Concentration level and inclusion criteria

Data exploration revealed that not all concentrations levels are provided in the same unit of measurement for all nanomaterials. For the analysis presented in this paper, only observations with concentration in $\mu g/cm^2$ will be analyzed. However, all units of measurement can be included for the controls, since all controls equal to dose 0. Table 1 in the supplementary appendix shows the number of the controls and the number of observations available for each nanomaterial that was included in the analysis. Figure 2a shows the DNA damage for Ag 16.7 nm with controls in all units of measurements while figure 2b shows the result for Ag 16.7 nm with controls in $\mu g/cm^2$ only. Figure 2c shows the DNA damage for NM-110 (ZnO 147 nm) for which the units are all in $\mu g/ml$ and therefore was not included in the analysis.

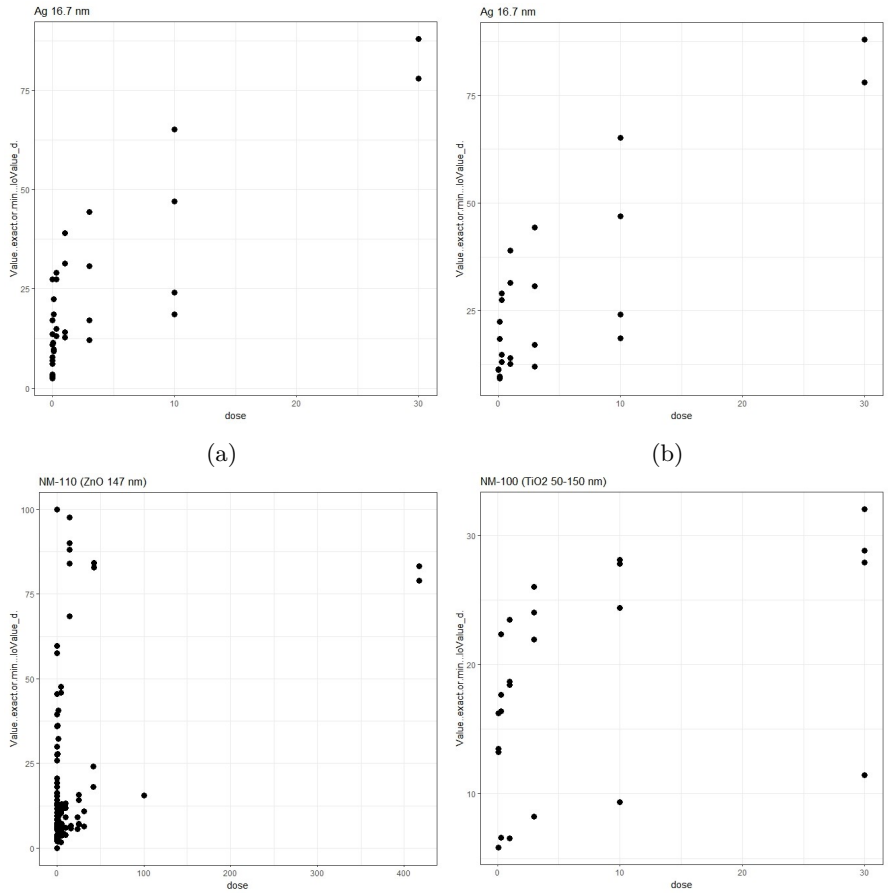


Figure 3: Dose-response plot of Ag 16.7 nm with controls measured in all units (a), and only in µg/cm2 (b), plot of NM-110 (ZnO 147 nm) (c), plot of NM-100 (TiO2 50-150 nm) (d)

Note that for several other nanomaterials, such as Polyoxyethylene Glycerol Trioleate, $BASO_4$ NM-220, NM-100 (TiO_2 50-150 nm), NM-212 Cerium (IV) Oxide precipitated, uncoated, NM-302 (Ag) and TiO_2, controls measurements are missing in the database and therefore were not included in the analysis (see also Figure 2d for an example of the NM-100 (TiO_2 50-150 nm) nanomaterial).

In addition, there are several nanomaterials with only 1 observation of control, such as Nanofibrillar cellulose 2-15 nm, Nanofibrillar cellulose 3-> nm, Nanofibrillar cellulose 5-10 nm, Nanofibrillar cellulose 7-20 nm and cellulose 30000 nm that were not included as well. In total, 26 nanomaterials are included in the analysis in which DNA damage is considered as the response variable.

2.2 Graphical exploration of the dose-response relationship

There are several other variables in the dataset that are of interest as they can be influence (or mask) the dose-response relationship. Figure 4 shows the *in vitro* data for NM 100 (Titanium Dioxide) and NM-110 (Zinc Oxide, uncoated) by the cell type used in the experiments. It shows that after differentiating the observations according to the cell type, increasing trends can be identified more clearly, especially for NM 100 (Titanium Dioxide).

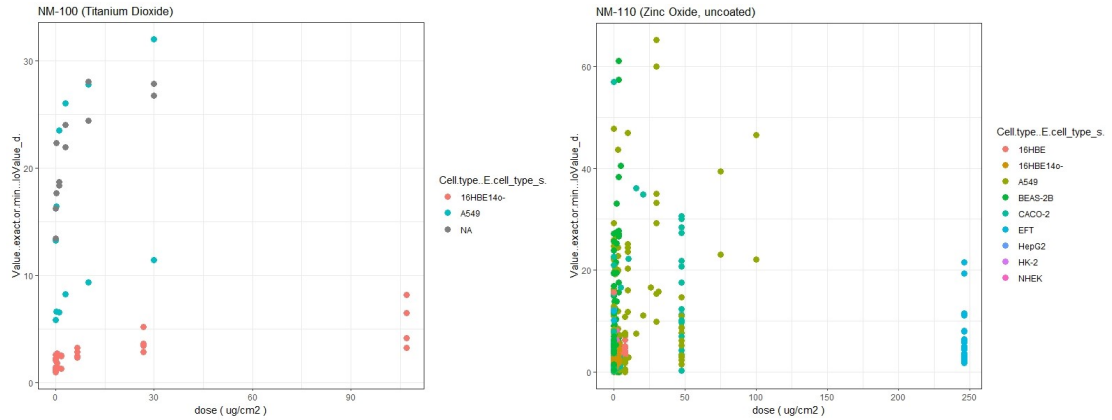


Figure 4: Dose-response plot of NM 100 (Titanium Dioxide) and NM-110 (Zinc Oxide, uncoated)

Figure 5 shows the *in vitro* data for the NM-100 nanomaterial by the method (assay) in which the experiment was conducted and by exposure time.

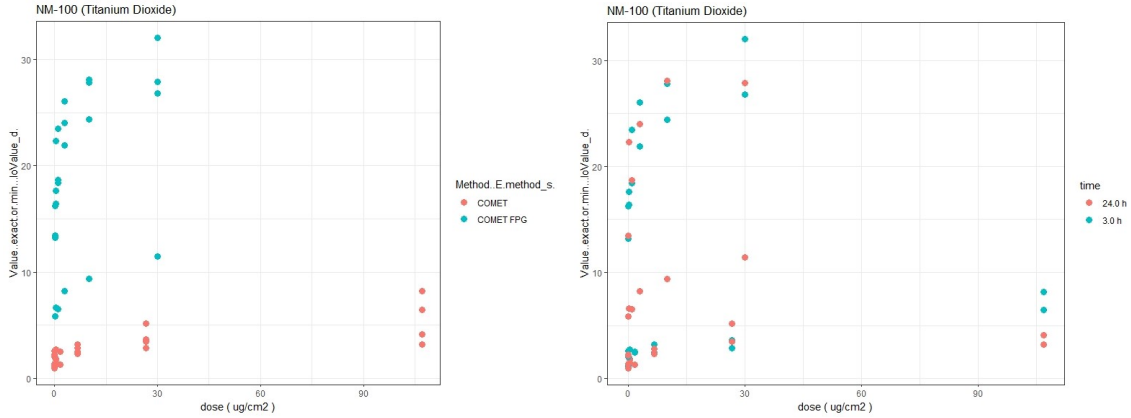


Figure 5: Dose-response plot of NM 100 (Titanium Dioxide)

3. Methodology

3.1 One-Way ANOVA under a simple order restriction.

Suppose that the response variable (DNA strand breaks) is denoted by Y and the concentration level by X_i, $i = 0, \ldots, K$ with X_0 being a control dose. Assume that there are in total $K + 1$ discrete concentrations, with $X_0 \leq X_1 \leq ... \leq X_K$ and n_i number of observations at concentration X_i. For a specific nanomaterial, the observed data can then be written in pairs as (x_i, y_{ij}), m=1,...,M. Let $\mu(x_i)$ be the mean at the ith dose level. Our aim is to test the homogeneity of the means against an order-restricted alternative. We consider a nanomaterial specific one way ANOVA model of the form.

$$Y_{ijm} = \mu(x_i) + \varepsilon_{ij}. \tag{1}$$

Here, as mentioned above Y_{ijm} denotes the DNA damage of the jth replicate in the ith dose level for the mth nanomaterial. The null hypothesis of no dose effect and the alternative hypothesis of positive dose effect (monotonically increasing means) are formulated as follows:

$$\begin{aligned} H_0 &: \mu(x_0) = \mu(x_1) = ... = \mu(x_K), \\ H_1 &: \mu(x_0) \leq \mu(x_1) \leq ... \leq \mu(x_K). \end{aligned} \tag{2}$$

Note that according to the type of the toxicity endpoint measured, the direction of the ordered mean is assumed to be known. For a decreasing dose-response trend the alternative hypothesis is given by

$$H_1 : \mu(x_0) \geq \mu(x_1) \geq ... \geq \mu(x_K). \tag{3}$$

3.2 Inference for order restricted alternative

There are several methods that can be applied to test the monotonic trend of the means (Lin et al., 2012). The methods of Williams, and of Marcus, are t-tests for order restricted inference, and are discussed in Section 3.2.1 while the likelihood ratio test is discussed in Section 3.2.2.

3.2.1 Williams and Marcus

For an experiment with n replicates of observation at each dose level, Williams test statistics (Williams, D. A., 1971) is given by:

$$t_i = \frac{\hat{\mu}^\star(x_i) - \bar{y}_0}{\sqrt{2s^2/n}}. \tag{4}$$

with $\hat{\mu}^\star(x_i)$ denotes the estimate of the isotonic mean at dose i, $\bar{y}_0$ the sample mean at dose 0 and s^2 the estimate of the variance (Lin et al., 2012). The isotonic mean at each dose level $\hat{\mu}^\star(x_i)$ are used in the step down procedure to determine the lowest dose for which a dose effect is detected (Williams, 1971). In case the number of observations is not equal at each dose level, the test statistics can be modified as follows:

$$t_i = \frac{\hat{\mu}^\star(x_i) - \bar{y}_0}{\sqrt{s^2/n_i + s^2/n_0}}. \tag{5}$$

Marcus modified Williams test statistics, by replacing $\bar{y}_0$ with the mean estimate at dose 0 under order restriction, $\hat{\mu}^\star(x_0)$ (Lin et al., 2012). Thus, the test statistics becomes

$$t_i = \frac{\hat{\mu}^\star(x_i) - \hat{\mu}^\star(x_0)}{\sqrt{s^2/n_i + s^2/n_0}}. \tag{6}$$

3.2.2 Likelihood ratio test

Another method that can be used to test the equality of the mean response under order restriction is through a Likelihood ratio test. This test can be used to detect a monotone trend but cannot give an indication in which dose(s) there is a difference. The test statistics is given by:

$$\Lambda_{01}^{\frac{2}{N}} = \frac{\hat{\sigma}^2_{H_1}}{\hat{\sigma}^2_{H_0}} = \frac{\sum_{ij}(y_{ijm} - \hat{\mu}_i^\star)^2}{\sum_{ij}(y_{ijm} - \hat{\mu})^2}, \tag{7}$$

with $\hat{\mu} = \sum_{ij} y_{ij} / \sum_i n_i$. It is the ratio between error variance under H_0 and error variance under H_1. The test statistics can also be written in term of

$$\bar{E}^2_{01} = 1 - \Lambda_{01}^{\frac{2}{N}} \tag{8}$$

The null hypothesis is rejected for a large value of $\bar{E}^2_{01}$ or for a small value of $\Lambda_{01}^{\frac{2}{N}}$. When the direction of the trend is unknown, the more likely direction can be chosen by comparing likelihood of the increasing and decreasing trend, the direction with the higher likelihood is selected.

3.2.3 Inference for high dimensional dose-response experiments

Due to the large number of nanomaterials for which the test for monotone trend is applied, adjustment for multiplicity should be considered. For the analysis presented in this paper, resampling based inference are used since the sample size of the experiments can be small and the assumption of the distribution of the

response might not be fulfilled (Lin et al., 2012). P-values can then be obtained through permutations, by recalculating the test statistics in each permutation. Suppose that permutation matrix $\mathbf{T}$ is

$$\mathbf{T} = \begin{pmatrix} t_{11} & t_{12} & ... & t_{1B} \\ t_{21} & t_{22} & ... & t_{2B} \\ \vdots & \vdots & \vdots & \vdots \\ t_{m1} & t_{m2} & ... & t_{mB} \end{pmatrix}, \tag{9}$$

with t_{mb} as the test statistics for the mth nanomaterial in the bth permutation b. The total number of the permutations is denoted by B. Let t_m be the test statistic for the mth nanomaterials, the row p-values can be calculated as:

$$P_m = \frac{\#(b : |t_{mb}| \geq |t_m|)}{B}. \tag{10}$$

Adjusting for multiple testing can be conducted by controlling the False Discovery Rate (FDR, Lin et al., 2012) or by controlling Family Wise Error Rate (FWER, Lin et al., 2012). An elaborate discussion about the procedures for multiplicity adjustment is given in the supplementary appendix of the paper (Section 1).

4. Application to the data

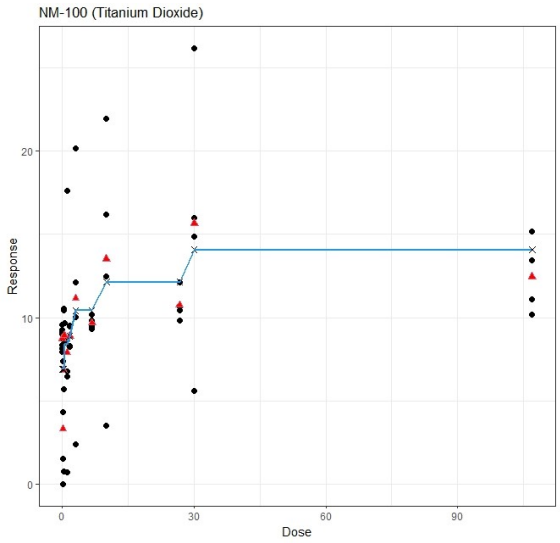


Figure 6: Dose-response plot of NM 100 (Titanium Dioxide) with isotonic means

4.1 Testing the monotonic trend of NM 100 (Titanium Dioxide)

For illustration, we present in this section the analysis for Titanium Dioxide. A pre-processing step in which the cell effect was removed from the data was applied (and described in Section 2 in the supplementary appendix). Figure 6 shows the dose-response plot for the pre-processed NM-100 (Titanium Dioxide) and the estimated isotonic mean at each dose-level.

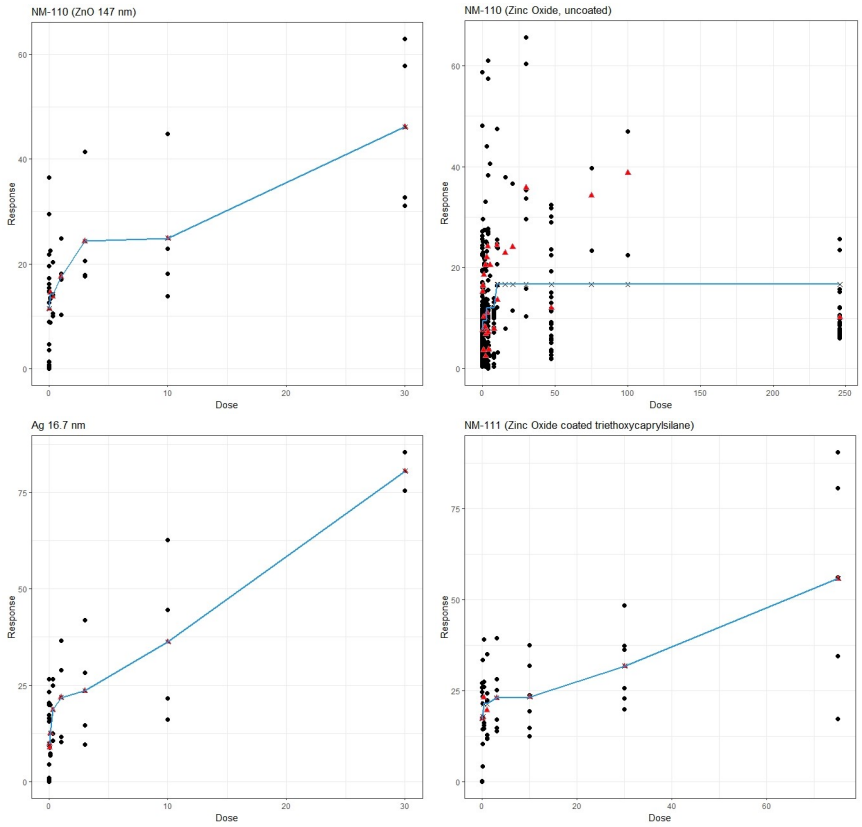
To test the monotonic dose-response relationship, the likelihood ratio test was applied. As shown in Table 1, the test statistics for NM-100 (Titanium Dioxide) is equal to 0.268 with p-value equals to 0.003 and therefore the null hypothesis, specified in Equation (2), of no dose effect is rejected.

Table 1: Test statistics and p-values for NM-100.

Nanomaterial	E2	p-value
NM-100 (Titanium Dioxide)	0.268	0.003

4.2 Testing the monotonic trends of all nanomaterials

The analysis presented in the previous section was applied to one NM, Titanium Dioxide, in this section we focus on the analysis of the entire dataset. Since 26 nanomaterials were tested simultaneously, the multiplicity correction, using the Benjamini and Hochberg (BH) procedure was applied. Table 2 presents the test statistics and p values and shows that, after adjusting for multiplicity, 13 nanomaterials are found to be have a significant dose-response relationship (see also Figure 6). Figure 7 shows the top 4 nanomaterials (NM-110 (ZnO 147 nm), NM-110 (Zinc Oxide, uncoated), Ag 16.7 nm and NM-111 (Zinc Oxide coated triethoxycaprylsilane)) that were found to be significant, the NM-202 (Synthetic Amorphous Silica PY-AB-03) with raw p-value equal to 0.046 and NM-203 (Synthetic Amorphous Silica PY-A-04) with raw p value equal to 0.080. Note that the later two are not significant after adjusting for multiplicity.



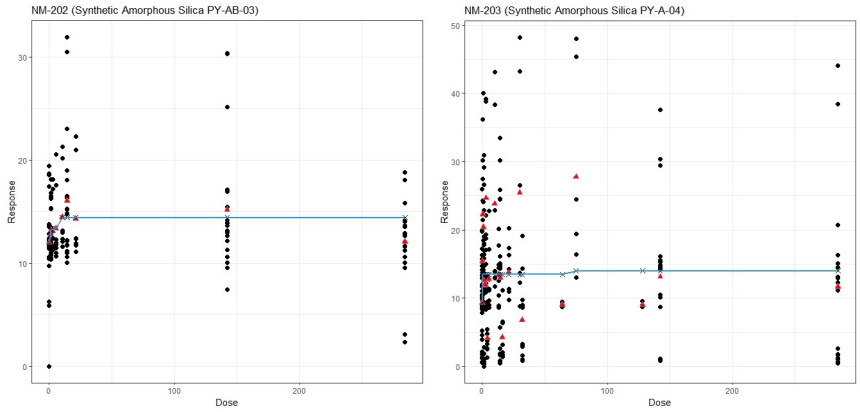


Figure 7: Isotonic mean plot of 6 nanomaterials

Table 2: Test statistics and p-values for all Nanomaterials in the dataset

Nanomaterial	E2	p-value	adj. p-value
NM-110 (ZnO 147 nm)	0.523	0.000	0.000
NM-110 (Zinc Oxide, uncoated)	0.115	0.000	0.000
Ag 16.7 nm	0.706	0.000	0.000
NM-111 (Zinc Oxide coated triethoxycaprylsilane)	0.432	0.000	0.000
BaSO4 25 nm	0.522	0.002	0.010
NM-100 (Titanium Dioxide)	0.268	0.003	0.011
NM-105 (Titanium Dioxide)	0.184	0.003	0.011
NM-401 (MWCNT 64.2 nm)	0.473	0.004	0.013
NM-200 (silica 18.3 nm)	0.244	0.006	0.017
NM-401 (Multi-walled carbon nanotubes)	0.046	0.007	0.018
NM-102 (Titanium Dioxide, anatase)	0.066	0.009	0.021
NM-400 (Multi-walled carbon nanotubes)	0.032	0.012	0.024
NM-200 (Synthetic Amorphous Silica PR-A-02)	0.052	0.012	0.024
NM-101 (TiO2 6 nm)	0.162	0.039	0.072
NM-212 (CeO2 33 nm)	0.131	0.047	0.076
NM-202 (Synthetic Amorphous Silica PY-AB-03)	0.044	0.046	0.076
NM-203 (Synthetic Amorphous Silica PY-A-04)	0.024	0.080	0.122
NM-101 (Titanium Dioxide)	0.072	0.122	0.176
NM-103 (TiO2 24.7 nm)	0.197	0.194	0.265
NM-104 (Titanium Dioxide)	0.026	0.282	0.357
NM-103 (Titanium Dioxide)	0.024	0.288	0.357
NM-201 (Synthetic Amorphous Silica PR-B-01)	0.005	0.550	0.650
MWCNT (Mitsui)	0.002	0.597	0.675
NM-402 (Multi-walled carbon nanotubes)	0.000	0.667	0.694
NM-403 (Multi-walled carbon nanotubes)	0.001	0.658	0.694
MWCNT (Cheap Tubes)	0.000	0.707	0.707

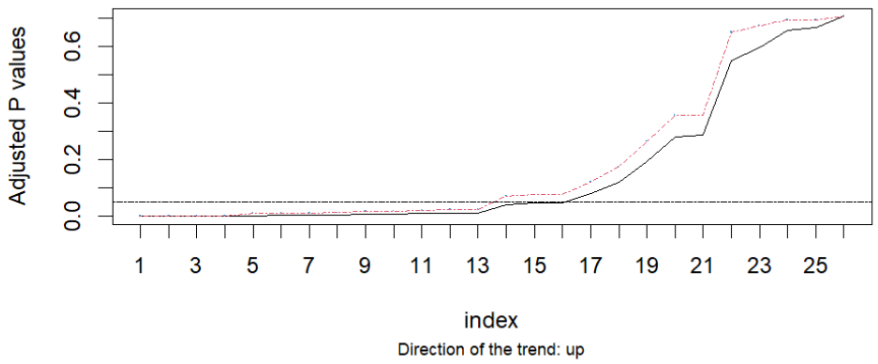


Figure 8: Adjusted p-values (using the BH-FDR procedures). Red line: adjusted p-values, Black line: raw p-values

5. Discussion

DNA strand breaks, which are an indicator of toxicity are assumed to monotonically increases as the concentration of the nanomaterial increases. To evaluate this trend in the dose-response relationship, a likelihood ratio test was applied on 26 nanomaterials in the dataset. Initially, there were 55 nanomaterials in the dataset, but we included only nanomaterials with concentration measured in $\mu g/cm^2$, excluded nanomaterials without control values and omitted concentration level with only 1 observation. In the end, there were only 26 nanomaterials left to be evaluated. According to the graphical exploration, variable cell type seemed to have an effect on the dose-response relationship. Therefore, prior to the testing, analysis of variance was used to adjust for the effect of variable cell type. There were several other variables that might also have an influence on the dose-response relationship, such as the exposure time, the method used in the experiment and the provider who performed the experiment. For future research, it might also be interesting to take into account these variables in the analysis. In addition to test the monotonic trend, dose-response modeling can be used to get more insight about the dose-response relationship, such as to identify the dose at which 50% of the effect is observed. Physico-chemical property of the nanomaterial, which may affect how the nanomaterial is taken up by the body is another factor that may also be interesting to explore.

Reference

eNanoMapper (n.d.). *Nanosafety projects data access*. Retrieved March 28, 2022, from https://search.data.enanomapper.net/projects/

NanoInformaTIX (2019). *NanoInformaTIX. Development and Implementation of a Sustainable Modelling Platform for NanoInformatics*. https://www.nanoinformatix.eu/

Gubala, V., Johnston, L. J., Krug, H. F., Moore, C. J., Ober, C. K., Schwenk, M., & Vert, M. (2018). Engineered nanomaterials and human health: Part 2. Applications and nanotoxicology (IUPAC Technical Report). Pure and Applied Chemistry, 90(8), 1325-1356.

Hochella, M. F., Mogk, D. W., Ranville, J., Allen, I. C., Luther, G. W., Marr, L. C., ... & Sahai, N. (2019). Natural, incidental, and engineered nanomaterials and their impacts on the Earth system. Science, 363(6434).

Hothorn, L. A. (2016). Statistics in toxicology using R. CRC Press.

Jeevanandam, J., Barhoum, A., Chan, Y. S., Dufresne, A., & Danquah, M. K. (2018). Review on nanoparticles and nanostructured materials: history, sources, toxicity and regulations. Beilstein journal of nanotechnology, 9(1), 1050-1074.

Jeliazkova, N., Chomenidis, C., Doganis, P., Fadeel, B., Grafström, R., Hardy, B., Hastings, J., Hegi, M., Jeliazkov, V., Kochev, N., Kohonen, P., Munteanu, C. R., Sarimveis, H., Smeets, B., Sopasakis, P., Tsiliki, G., Vorgrimmler, D., & Willighagen, E. (2015). The eNanoMapper database for nanomaterial safety information. Beilstein journal of nanotechnology, 6, 1609–1634. https://doi.org/10.3762/bjnano.6.165/
Kabir, E., Kumar, V., Kim, K. H., Yip, A. C., & Sohn, J. R. (2018). Environmental impacts of nanomaterials. Journal of Environmental Management, 225, 261-271.

Lin, D., Shkedy, Z., Yekutieli, D., Amaratunga, D., & Bijnens, L. (Eds.). (2012). Modeling dose-response microarray data in early drug development experiments using R: order-restricted analysis of microarray data. Springer Science & Business Media.

Otava, M., Sengupta, R., Shkedy, Z., Lin, D., Pramana, S., Verbeke, T., ... & Klinglmueller, F. (2017). IsoGeneGUI: Multiple approaches for dose-response analysis of microarray data using R. The R Journal, 9(1), 14-26.

Pramanik, P., Krishnan, P., Maity, A., Mridha, N., Mukherjee, A., & Rai, V. (2020). Application of Nanotechnology in Agriculture. In Environmental Nanotechnology Volume 4 (pp. 317-348). Springer, Cham.

Van Miert, S., Creylman, J., & Verheyen, G. R. (2019). Mining a Nanoparticle Dataset, Compiled Within the MODENA-COST Action. International Journal of Quantitative Structure-Property Relationships (IJQSPR), 4(1), 1-17.

Williams, D. A. (1971). A test for differences between treatment means when several dose levels are compared with a zero dose control. Biometrics, 103-117.

Yokel, R. A., & MacPhail, R. C. (2011). Engineered nanomaterials: exposures, hazards, and risk prevention. Journal of Occupational Medicine and Toxicology, 6(1), 1-27.