**Research Paper**
Academic Platform Journal of Engineering and Smart Systems (APJESS) 10(3), 168-173, 2022
ISSN: 2822-2385
Received: 26-Jun-2022 Accepted: 12-Sep-2022
https://doi.org/10.21541/apjess.1135939

# Modelling Genotoxic Effects of Metal Oxide Nanoparticles using QSAR Approach

*[1]Ceyda ÖKSEL KARAKUŞ

[1]Department of Bioengineering, Izmir Institute of Technology, İzmir, Turkey, ceydaoksel@gmail.com

## Abstract

We investigated the application of structure-activity relationship approaches to underpin structural properties that potentially control the genotoxic potential of 9 different metal oxide nanoparticles (CuO, ZnO, NiO, $SiO_2$, $TiO_2$, $CeO_2$, $Fe_2O_3$, $Fe_3O_4$ and $Co_3O_4$). In particular, we compiled a pool of quantum-mechanical, experimental and periodic table-driven descriptors and explored their distinctive contribution to the measured activity (genotoxicity). We first employed a clustered heatmap and parallel coordinates plot for visual exploration of the clusters and outliers of the data and finding corresponding responsible physicochemical descriptors. We then investigated the strength (and direction) of the relationship among descriptors and between descriptors and genotoxicity using similarity metrics. By using orthogonal projections to latent structures (OPLS), we were able to quantify the relative contribution of each descriptor to the genotoxicity of metal oxide nanoparticles. Our results suggested that zeta potential, the ratio of core electrons to valence electrons, Fermi energy and electronegativity were significant predictors of genotoxicity. Such computer-assisted approaches hold considerable promise for maximizing the use of accumulated data in nanotoxicology, prioritizing nanoparticles for further testing and filling data gaps required for hazard assessment processes.

***Keywords:*** Genotoxicity; Nanoparticles; Nanotoxicology; Computational toxicology; QSAR

## 1. INTRODUCTION

Ongoing globalization of nano-sciences has not only made exposure to nano-enabled materials inevitable, but also raised new challenges in the safety and regulatory domains. In particular, unique physicochemical properties of metal oxide (MO) nanoparticles (NPs) have led to an ongoing increase in their manufacture and commercial use. Studies to date have shown that MONPs are capable of inducing cytotoxic [1], genotoxic/oxidative [2], neurotoxic [3, 4] and inflammatory effects [5]. Despite recent collaborative research efforts to tackle safety issues linked to MONPs, a consensus on how to properly assess their health risks has not been reached yet. Advances in molecular biology and information science have led to a paradigm shift in hazard assessment of both traditional and advanced chemicals, from heavily relying on animal-based tests, towards developing non-testing alternatives. The term “non-testing” in the context of hazard assessment refers to the use of computational technologies and to make sound risk management decisions regarding the safe manufacturing and use of products at lower costs and in less time along the path from design to commercial manufacturing and use. Non-testing approaches hold considerable promise for making better use of accumulated hazard data and contributing to the needs of the safety assessment of new chemicals at early design stages.

Non-testing methods are generally placed into three broad categories: 1) grouping and read-across approaches, 2) structure-activity relationship (SAR) methods and 3) expert systems. While each of these methods differs in advanced settings with regard to the type of input data, methodology and level of complexity involved, they all depend on the same assumption that structurally similar compounds behave similarly in biological systems. In the context of chemical safety assessment, machine learning and statistical approaches correlating activity with the structure are widely used for categorization of chemicals based on hazard potential and prioritization for further testing, identification of hazard and hazard-related physicochemical properties and minimization of hazard by modifying toxicity-related properties (so-called *safe-by-design*). While the use of computational approaches in the safety assessment of conventional (drug-like) compounds is common practice, especially in pharmaceutical sciences, their application to NPs is still in its infancy and requires further investigation.

* Corresponding Author

Here, we related the genotoxicity of MONPs that are commonly used in commercial applications to their physicochemical properties using SAR approaches. In particular, the relationship between descriptors encoding physicochemical properties and their contribution to the genotoxicity of MONPs was visually explored by high-dimensional data visualization tools and modelled by a modified version of the partial least squares regression technique. A Pearson's correlation matrix is employed to quantify the direction and strength of the relationship among descriptors and between descriptors and genotoxicity.

## 2. MATERIALS AND METHODS

### 2.1. Toxicity Data

Data on the genotoxic potential of 9 different MONPs (CuO, ZnO, NiO, $SiO_2$, $TiO_2$, $CeO_2$, $Fe_2O_3$, $Fe_3O_4$ and $Co_3O_4$) have been collected from the literature [6-8]. The collected toxicity data involve NP-induced DNA damage in lung epithelial cells (A549) measured by the alkaline version of the Comet assay following 4-hr exposure to 20 μg/cm$^2$ and 40 μg/cm$^2$ of MONPs.

### 2.2. Descriptors

Descriptor data used in SAR analysis are given in Table 1.

**Table 1.** Experimentally-measured, quantum mechanically-computed and periodic table-based descriptors used in this study

| | No | Descriptor | Explanation |
|---|---|---|---|
| Experimental | EX1 | Size_TEM | Mean particle size |
| | EX2 | Size_DLS | Hydrodynamic particle size |
| | EX3 | Zeta | Zeta potential |
| | EX4 | ROS | Reactive Oxygen Species |
| Quantum Mechanical | QM1 | F_energy | Fermi energy |
| | QM2 | HOMO | Highest occupied molecular |
| | QM3 | LUMO | Lowest unoccupied energy |
| | QM4 | B_gap | HOMO-LUMO band gap |
| | QM5 | Hardness | Chemical hardness |
| | QM6 | C_pot | Chemical potential |
| | QM7 | ΔH_f | Enthalpy of formation |
| | QM8 | χ | Electronegativity |
| Periodic Table-based | PT1 | MW | Molecular weight, g/mol |
| | PT2 | No_Me | Number of metals |
| | PT3 | No_Ox | Number of oxygen atoms |
| | PT4 | EN_Me | Electronegativity of metal |
| | PT5 | T_EN_Me | Total metal electronegativity |
| | PT6 | T_EN/No_Ox | Electronegativity per oxygen |
| | PT7 | Ox_St | Oxidation state |
| | PT8 | Z_Me | Atomic number |
| | PT9 | Zv_Me | Valence electrons of metal |
| | PT10 | PT_Me | Periodic table number of metal |
| | PT11 | (Z_Zv)/Zv | Core electrons/valence |
| | PT12 | V_Me | Valency of metal |

#### 2.2.1. Experimental Descriptors

A set of four experimental descriptors including mean particle size measured by Transmission Electron Microscopy (TEM), hydrodynamic diameter and zeta potential measured by Dynamic Light Scattering (DLS) and reactive oxygen species (ROS) formation capacity was collected from literature [6-8] and used as input parameters in SAR analysis.

#### 2.2.2. Quantum-Mechanical Descriptors

A set of 8 quantum mechanical descriptors derived based on density functional theory (DFT) including Fermi energy, highest occupied molecular orbitals (HOMO), lowest unoccupied energy orbitals (LUMO), band gap, hardness, chemical potential, enthalpy of formation, and electronegativity was collected from the literature [9].

#### 2.2.3. Periodic Table-Based Descriptors

A set of 12 descriptors including molecular weight, number of metals and oxygen atoms, electronegativity of metal, total metal electronegativity per oxygen atoms, oxidation state, atomic number, periodic table number of metal, the ratio of the number of core electrons to the number of valence electrons and valency of metal was calculated based on the molecular formula and the periodic table of the elements.

### 2.3. Data Scaling

Prior to data exploration and modelling, the z-score transformation was carried out for normalizing each value in the dataset using the following formula:

$$z - score = \frac{actual\ value - \mu}{\sigma}$$

where μ is the mean and σ is the standard deviation.

### 2.4. Exploratory Analysis

For the visual exploration of multivariate descriptor data in relation to toxicity potential, two different techniques, parallel coordinates and heatmaps, were employed. Pearson correlation matrix was constructed to determine inter-correlated descriptors and how strongly each of these descriptors is related to toxicity. All analyses were carried out in R Version 4.2.0 [10] using the packages "pheatmap", "Hmisc", "corrplot", and "PerformanceAnalytics".

### 2.5. Modelling

Partial least squares (PLS) is a modelling technique that combines dimensionality reduction through principal component analysis (PCA) and regression [11]. While unsupervised PCA is mainly focusing on finding the compressed representation of predictor variables only, in supervised PLS regression, the correlation between the predictor and outcome variables is taken into account to compute principal components which are then used to fit a regression model. One of the most recent extensions of PLS is called orthogonal projections to latent structures (OPLS) which decomposes the PLS solution into predictive

(correlated with the response) and orthogonal (uncorrelated with the response) components [12]. To model the relationship between descriptors and the genotoxicity potential of MONPS, OPLS regression was performed on the scaled data using the SIMCA software version 17.

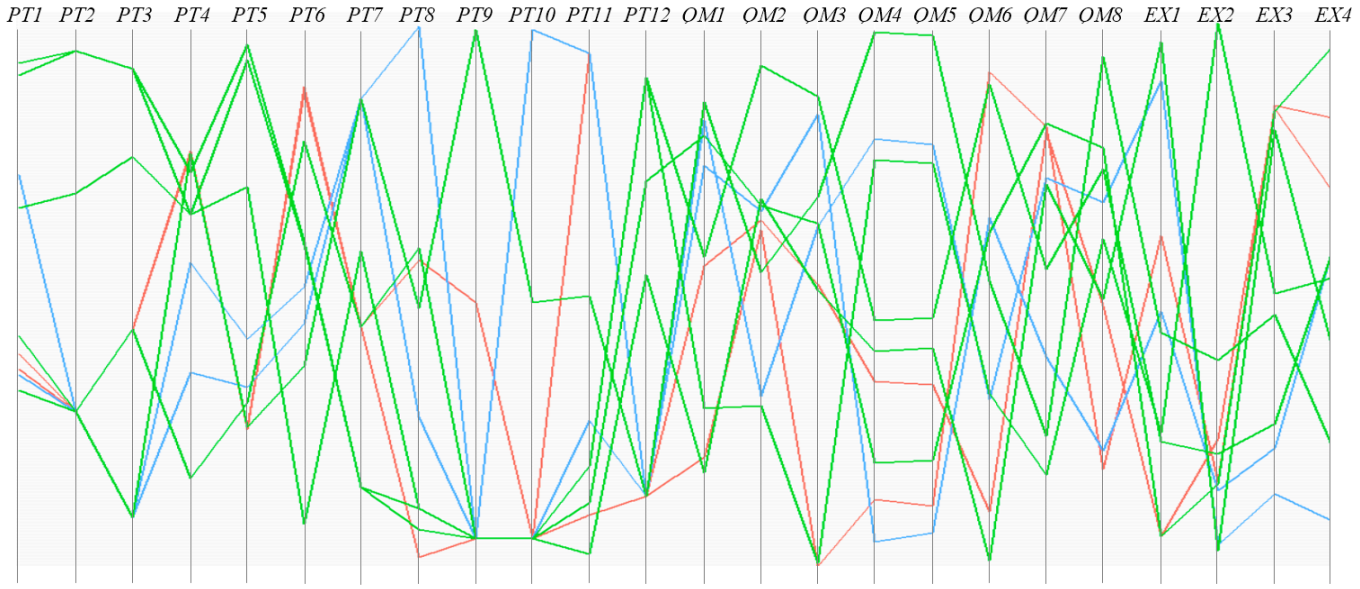


**Figure 1.** Periodic table (PT) descriptors (PT1-12), Quantum Mechanical (QM) descriptors (QM1-8) and Experimental (EX) descriptors (EX1-4) of 9 MoNPs coloured according to high (red), medium (blue) and low (green) values of genotoxicity at 40 μg/cm$^2$

## 3. RESULTS

For the purpose of data exploration, three different techniques including parallel coordinates, heat maps and Pearson's correlation matrix were used. In parallel coordinates, the points used in Euclidean space are represented as a series of lines passing through parallel axes where each variable is represented by one parallel axis. Parallel coordinate visualization of the descriptor data is given in Figure 1. The results associated with high, medium and low genotoxicity are highlighted in red, blue and green, respectively. The parallel coordinate plot shows that electronegativity (QM8) is inversely related to genotoxicity while higher values of zeta potential (EX3) lead to a higher genotoxicity level. Similar observations can be made from Pearson's correlation matrix given in Figure 2.

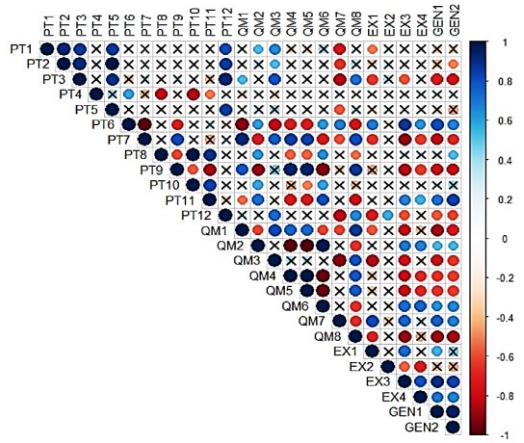


**Figure 2**. Pearson's correlation matrix displaying correlations between descriptors and genotoxicity [from –1 (blue) to +1 (red)]. Color intensity and circle size are proportional to the strength of correlation.

In Figure 2, positive and negative correlations are displayed in blue and red, respectively, while x indicates statistically non-significant correlations based on p-value of 0.05. GEN1 and GEN2 refer to NP-induced DNA damage in A549 following 4-hr exposure to 20 μg/cm$^2$ and 40 μg/cm$^2$ of MONPs. It is clear from Figure 2 that number of Oxygen atoms (PT3), valence electrons of metal (PT9), Fermi energy (QM1) and electronegativity (QM8) are negatively correlated with DNA damage while the ratio of core electrons to valence electrons (PT11) and zeta potential (EXP3) are positively correlated with DNA damage. We then combined heat map with clustering to prioritize MONPs based on genotoxic potential and to demonstrate the physicochemical differences between the ones belonging to different genotoxicity clusters (Figure 3).

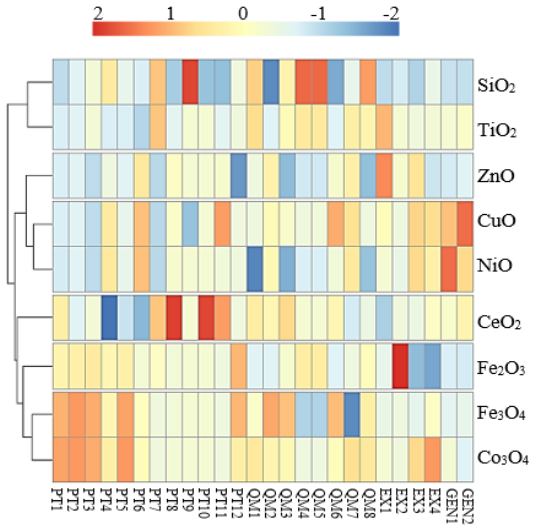


**Figure 3.** Clustered heat map displaying auto-scaled descriptor and genotoxicity values of 9 MONPs

Several observations can be made from the cluster heat map given in Figure 3. In particular, two MONPs, CuO and NiO are clearly distinguished from the rest due to their relatively high genotoxicity potential. Another interesting finding is that $SiO_2$ which has the lowest genotoxicity at both measured doses also has the most extreme (e.g. the lowest or the highest) values in 14 descriptors (out of 24), suggesting the great influence of physicochemical properties on toxicity.

As a next step, OPLS has been carried out to model the relationship between 24 descriptors and 2 genotoxicity results simultaneously. The $R^2$ values reflecting the model fit are found to be 0.83 and 0.85 for genotoxicity measured at 20 and 40 μg/cm$^2$, respectively. The resulting score scatter, loadings and variable importance plots are given in Figures 4–6.

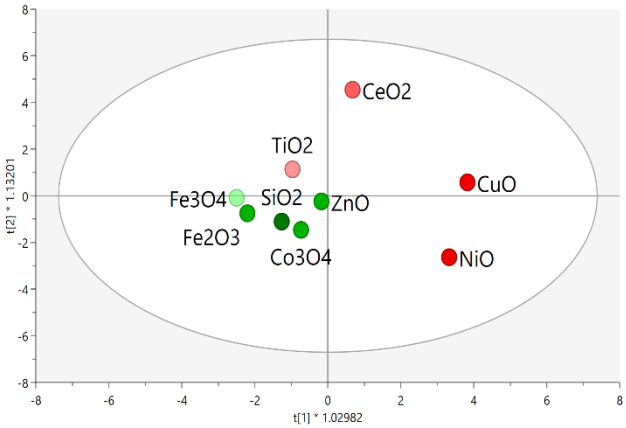


**Figure 4.** OPLS t1/t2 score plot showing the relationship between MONPs (red and green indicate higher and lower values of genotoxicity, respectively)

As can be seen from Figure 4, three particular MONPs with relatively higher genotoxicity potential, CuO, NiO and $CeO_2$, are clearly separated from the main cluster formed of MONPs with lower-genotoxicity values. The OPLS loadings given in Figure 5 display how the x-variables are combined to form the OPLS-score plot given in Figure 4.

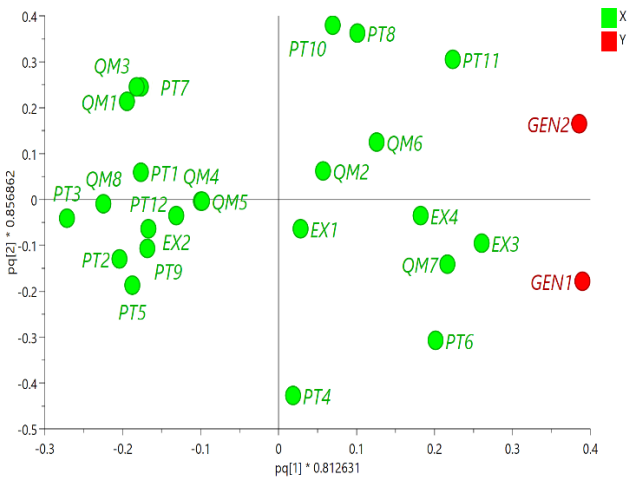


**Figure 5.** OPLS loadings plot showing the relationship between 24 descriptors (x variables) and 2 genotoxicity responses (y variables)

By looking at the loading plots given in Figure 5, we can identify the descriptors contributing to the positioning and separation of MONPs shown in Figure 4. For the ease of interpretation, one can imagine a straight line passing through the response variable (GEN1 or GEN2) and the origin of the plot. When the descriptors (x variables) are perpendicularly projected onto this imaginary line, their distance from the origin signifies their contribution to the response variable. Based on this, it can be concluded that mean particle size measured by TEM (EX1) has a near-zero contribution to genotoxicity while zeta potential (EX3), number of oxygen atoms (PT3), the ratio of core electrons to valence electrons (PT11) and electronegativity (QM8) have the largest impact on genotoxic potential.

The variable importance in projection (VIP) plot showing the relative contribution of each descriptor to genotoxicity is given in Figure 6. The variables with the highest VIP scores are the most influential parameters for genotoxicity.

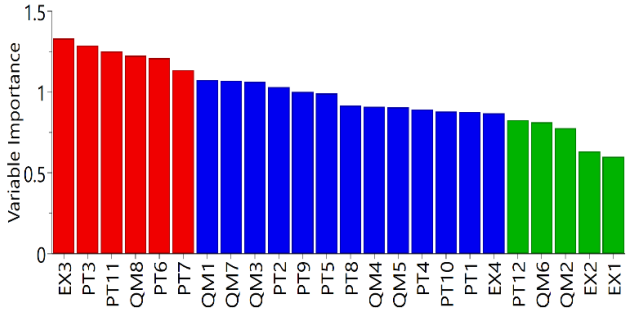


**Figure 6.** Variable importance in projection (VIP) plot summarizing each variable's contribution to the OPLS model

## 4. DISCUSSION

Using a combination of multidimensional data visualization techniques, similarity metrics, clustering and regression-based models, we identified intrinsic physicochemical properties that are seemingly related to the genotoxicity potential of MONPs. To ensure the collection of homogenous genotoxicity data, we restricted our selection to a specific cell line (lung epithelial cells), assay (alkaline Comet assay), NP class (metal oxides), exposure duration (4 hours) and concentration (20 and 40 μg/cm$^2$). We combined literature collected experimental [6-8] and theoretical [9] descriptor data with simple periodic-table driven descriptors to encode the physicochemical properties of MONPs.

Our results suggested that NiO and CuO, induced the highest level of DNA damage in the human lung epithelial cells, followed by $CeO_2$ and $TiO_2$. Indeed, the high genotoxic potential of both NPs has been frequently reported in the literature [13-18]. The next step was to explore the physicochemical drivers of the observed high and low genotoxicity. Out of 24 experimentally-measured and theoretically calculated descriptors, four particular ones, zeta potential, number of oxygen atoms, the ratio of core electrons to valence electrons and electronegativity were repeatedly found to be correlated with the observed genotoxicity. In particular, highly positive (>30 mV) values of low molecular weight (<80 g/mol) MONPs led to high

genotoxicity (>20% DNA damage compared to the control). The close correlation between positive surface charge and high toxicity is generally attributed to the increased capacity of positively-charged MONPs to interact with the negatively-charged cell membrane leading to higher levels of cellular uptake [19]. Despite its high zeta potential, $Co_3O_4$ did not induce any DNA damage (potentially) due to its high molecular weight (241 g/mol) and/or the high number of oxygen atoms.

A reverse relationship between electronegativity and genotoxic potential was observed. This could be attributed to the metal atoms with low electronegativity triggering an increase in the ionic concentration inside the endosome membrane (so-called proton sponge effect) and consequently leading to a swelling structure due to osmotic pressure, lysosome damage and cell death [20]. Another significant parameter contributing to the genotoxicity of MONPs was the ratio of the number of core electrons to the number of valence electrons. This trend was also observed by Kar, et al. [21] who reported that a higher value of the core and valence electron ratio resulted in higher levels of toxicity to E.coli.

## 5. CONCLUSION

The alternative non-testing methods and approaches under EU's REACH (Registration, Evaluation, Authorization and Restriction of Chemicals) regulation refer to the use of data-driven models for the purpose of NP labelling, categorization and prediction of toxicity, environmental fate and ecotoxicity. Ultimately, they are expected to support decision-making processes in the context of safety assessment by exploring data-driven estimates for occupational, consumer and environmental risks associated with the manufacture and use of NPs. In particular, non-testing methods hold great promise for maximizing the use of experimental toxicity data and contributing to the needs of the safety assessment of new materials at early design stages. While these techniques are not yet able to replace animal testing in nanotoxicology, they are capable of supporting 3R (Replacement, Reduction and Refinement) principles by minimizing the number of animals used in scientific procedures. Despite the strong industrial need and growing scientific and regulatory interest in the use of non-testing methods for prioritizing NPs according to their toxic potential and understanding structure-related root causes of their observed toxicity, the practical application and regulatory acceptance of computer-enabled material categorization and toxicity predictions have been limited so far. Given the diverse advantages that non-testing approaches offer to nanotoxicology, their use in regulatory context is expected to grow in the coming years.

**Author contributions:** Concept – C.O.K.; Data Collection &/or Processing – C.O.K.; Literature Search – C.O.K.; Writing – C.O.K.
**Conflict of Interest:** No conflict of interest was declared by the authors.
**Financial Disclosure:** The authors declared that this study has received no financial support.

## 6. REFERENCES

[1] X. Hu, S. Cook, P. Wang, and H.-m. Hwang, "In vitro evaluation of cytotoxicity of engineered metal oxide nanoparticles," Science of the Total Environment, vol. 407, no. 8, pp. 3070-3072, 2009.

[2] N. Golbamaki et al., "Genotoxicity of metal oxide nanomaterials: review of recent data and discussion of possible mechanisms," Nanoscale, vol. 7, no. 6, pp. 2154-2198, 2015.

[3] S. Singh, "Zinc oxide nanoparticles impacts: cytotoxicity, genotoxicity, developmental toxicity, and neurotoxicity," Toxicology mechanisms and methods, vol. 29, no. 4, pp. 300-311, 2019.

[4] B. Song, Y. Zhang, J. Liu, X. Feng, T. Zhou, and L. Shao, "Is neurotoxicity of metallic nanoparticles the cascades of oxidative stress?," Nanoscale research letters, vol. 11, no. 1, pp. 1-11, 2016.

[5] W.-S. Cho et al., "Metal oxide nanoparticles induce unique inflammatory footprints in the lung: important implications for nanoparticle testing," Environmental health perspectives, vol. 118, no. 12, pp. 1699-1706, 2010.

[6] H. L. Karlsson, J. Gustafsson, P. Cronholm, and L. Möller, "Size-dependent toxicity of metal oxide particles—a comparison between nano-and micrometer size," Toxicology letters, vol. 188, no. 2, pp. 112-118, 2009.

[7] J. Kain, H. Karlsson, and L. Möller, "DNA damage induced by micro-and nanoparticles—interaction with FPG influences the detection of DNA oxidation in the comet assay," Mutagenesis, vol. 27, no. 4, pp. 491-500, 2012.

[8] H. L. Karlsson, P. Cronholm, J. Gustafsson, and L. Moller, "Copper oxide nanoparticles are highly toxic: a comparison between metal oxide nanoparticles and carbon nanotubes," Chemical research in toxicology, vol. 21, no. 9, pp. 1726-1732, 2008.

[9] M. Thwala et al., "Using the Isalos platform to develop a (Q) SAR model that predicts metal oxide toxicity utilizing facet-based electronic, image analysis-based, and periodic table derived properties as descriptors," Structural Chemistry, vol. 33, no. 2, pp. 527-538, 2022.

[10] R. C. Team. "R: A Language and Environment for Statistical Computing." (accessed.

[11] H. Wold, "Soft modelling by latent variables: the non-linear iterative partial least squares (NIPALS) approach," Journal of Applied Probability, vol. 12, no. S1, pp. 117-142, 1975.

[12] J. Trygg and S. Wold, "Orthogonal projections to latent structures (O-PLS)," Journal of Chemometrics: A Journal of the Chemometrics Society, vol. 16, no. 3, pp. 119-128, 2002.

[13] M. Ahamed, M. A. Siddiqui, M. J. Akhtar, I. Ahmad, A. B. Pant, and H. A. Alhadlaq, "Genotoxic potential of copper oxide nanoparticles in human lung epithelial

cells," Biochemical and biophysical research communications, vol. 396, no. 2, pp. 578-583, 2010.

[14] M. J. Akhtar, S. Kumar, H. A. Alhadlaq, S. A. Alrokayan, K. M. Abu-Salah, and M. Ahamed, "Dose-dependent genotoxicity of copper oxide nanoparticles stimulated by reactive oxygen species in human lung epithelial cells," Toxicology and industrial health, vol. 32, no. 5, pp. 809-821, 2016.

[15] A. Semisch, J. Ohle, B. Witt, and A. Hartwig, "Cytotoxicity and genotoxicity of nano-and microparticulate copper oxide: role of solubility and intracellular bioavailability," Particle and fibre toxicology, vol. 11, no. 1, pp. 1-16, 2014.

[16] L. Capasso, M. Camatini, and M. Gualtieri, "Nickel oxide nanoparticles induce inflammation and genotoxic effect in lung epithelial cells," Toxicology letters, vol. 226, no. 1, pp. 28-34, 2014.

[17] M. Ahamed, D. Ali, H. A. Alhadlaq, and M. J. Akhtar, "Nickel oxide nanoparticles exert cytotoxicity via oxidative stress and induce apoptotic response in human liver cells (HepG2)," Chemosphere, vol. 93, no. 10, pp. 2514-2522, 2013.

[18] R. F. De Carli et al., "Evaluation of the genotoxic properties of nickel oxide nanoparticles in vitro and in vivo," Mutation Research/Genetic Toxicology and Environmental Mutagenesis, vol. 836, pp. 47-53, 2018.

[19] M. Weiss et al., "Density of surface charge is a more predictive factor of the toxicity of cationic carbon nanoparticles than zeta potential," Journal of nanobiotechnology, vol. 19, no. 1, pp. 1-19, 2021.

[20] Y. Huang et al., "Use of dissociation degree in lysosomes to predict metal oxide nanoparticle toxicity in immune cells: Machine learning boosts nano-safety assessment," Environment International, vol. 164, p. 107258, 2022.

[21] S. Kar, K. Pathakoti, P. B. Tchounwou, D. Leszczynska, and J. Leszczynski, "Evaluating the cytotoxicity of a large pool of metal oxide nanoparticles to Escherichia coli: Mechanistic understanding through In Vitro and In Silico studies," Chemosphere, vol. 264, p. 128428, 2021.