



LJMU Research Online

Gajewicz, A, Cronin, MTD, Rasulev, B, Leszczynski, J and Puzyn, T

Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: nano-read-across

<http://researchonline.ljmu.ac.uk/id/eprint/2932/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Gajewicz, A, Cronin, MTD, Rasulev, B, Leszczynski, J and Puzyn, T (2015) Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: nano-read-across. NANOTECHNOLOGY. 26 (1). pp. 1-9. ISSN 0957-4484

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

Novel approach for efficient predictions properties of large pool of nanomaterials based on limited set of species: Nano-read-across

Journal:	<i>Nanotechnology</i>
Manuscript ID:	Draft
Manuscript Type:	Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Gajewicz, Agnieszka; University of Gdansk, Faculty of Chemistry Cronin, Mark; School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Rasulev, Bakhtiyor; Interdisciplinary Nanotoxicity Center, Department of Chemistry and Biochemistry, Jackson State University, Leszczynski, Jerzy; Interdisciplinary Nanotoxicity Center, Department of Chemistry and Biochemistry, Jackson State University, Puzyn, Tomasz; University of Gdansk, Faculty of Chemistry
Article Keywords:	nanoparticles, nano-QSAR, nano-read-across, cytotoxicity
Abstract:	Creating suitable chemical categories and developing read-across methods, supported by quantum-mechanical calculations, can be an effective solution solving key problems related to current scarcity of data on the toxicity of various nanoparticles. This study has demonstrated that by applying a nano-read-across, the cytotoxicity of nano-sized metal oxides could be estimated with a similar level of accuracy as provided by nano-QSAR model(s). The method presented is a suitable computational tool for the preliminary hazard assessment of nanomaterials. It also could be used for the identification of nanomaterials that may pose potential negative impact to the human health and the environment. Such approaches are especially necessary, when there is paucity of relevant and reliable data points to develop and validate nano-QSAR models.

1
2
3
4 **Novel approach for efficient predictions properties of large pool**
5
6
7 **of nanomaterials based on limited set of species: Nano-read-**
8
9 **across**
10
11
12
13
14
15

16 *Agnieszka Gajewicz¹, Mark T.D Cronin², Bakhtiyor Rasulev³, Jerzy Leszczynski³ and*
17
18 *Tomasz Puzyn^{1*}*
19
20
21
22

23
24 ¹*Laboratory of Environmental Chemometrics, Institute for Environmental and Human Health*
25 *Protection, Faculty of Chemistry, University of Gdansk, Gdansk, Poland*
26

27
28 ²*School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom*
29 *Street, Liverpool L3 3AF, UK*
30

31
32 ³*Interdisciplinary Nanotoxicity Center, Department of Chemistry and Biochemistry, Jackson*
33 *State University, Jackson MS, USA;*
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52

53
54 ***Correspondence:** T. Puzyn, Laboratory of Environmental Chemometrics, Department of
55 Chemistry, University of Gdansk, Wita Stwosza 63, 80-308 Gdansk, Poland,

56
57 phone: (+48 58) 523 52 48; fax: (+48 58) 523 50 12; e-mail: t.puzyn@qsar.eu.org
58
59
60

Abstract

Creating suitable chemical categories and developing read-across methods, supported by quantum-mechanical calculations, can be an effective solution solving key problems related to current scarcity of data on the toxicity of various nanoparticles. This study has demonstrated that by applying a nano-read-across, the cytotoxicity of nano-sized metal oxides could be estimated with a similar level of accuracy as provided by nano-QSAR model(s). The method presented is a suitable computational tool for the preliminary hazard assessment of nanomaterials. It also could be used for the identification of nanomaterials that may pose potential negative impact to the human health and the environment. Such approaches are especially necessary, when there is paucity of relevant and reliable data points to develop and validate nano-QSAR models.

Keywords: nanoparticles, nano-QSAR, nano-read-across, cytotoxicity

1. Introduction

There is a large number of engineered nanomaterials of varying structure as well as those of the same chemical formula that differs only in terms of their physico-chemical properties. The differences in the physico-chemical properties of nanoparticles often determine the changes in their hazardous properties [1]. Unfortunately, it is still largely unknown, which properties determine and/or influence the toxicity of particular nanoparticles [2-4]. Therefore, effort should be placed into defining and developing methods for characterization, exposure, engineering control, potential toxicity, fate and transport, and life cycle assessment of engineered nanoparticles [5]. This leads to an increase in testing, and thus animal use, when traditional human health and the environment hazard assessments are conducted for all different nanomaterials. According to the recent report of the European Commission Joint Research Centre approximately 3.9 million additional test animals could potentially be used as a consequence of the introduction of REACH (Registration, Evaluation, Authorisation and Restriction of CHemicals) program [6] if the use of alternative methods would not be available. However, if these techniques would be applied in a maximal extend, a reduction in animal use could be obtained with potential savings of 1.9 million test animals [7]. Therefore, with regard to the ethical aspects, and following the recommendations by the REACH system and new European Directive 2010/63/EU [8], testing should be performed by means other than vertebrate animal tests, whenever possible. This requirement specifically encourages the use of alternative approaches, for example, *in vitro* methods or qualitative or quantitative structure-activity and structure-property relationship models

1
2
3
4 (QSAR/QSPR), or from information arising from the grouping of structurally related
5
6 substances (chemical category formation leading to read-across) [6].
7
8

9
10 The concept of using (Q)SAR methodologies in the risk assessment of
11
12 engineered nanomaterials has been discussed extensively for at least five years by
13
14 many national and international bodies (e.g., World Health Organization (WHO),
15
16 Organization for Economic Co-operation and Development (OECD), European
17
18 Commission (REACH, NanoSafety Cluster), European Food Safety Authority
19
20 (EFSA), Scientific Committee on Consumer Products (SCCP), Environmental
21
22 Protection Agency (EPA), etc.). (Q)SAR/QSPR approaches are based on defining
23
24 mathematical dependencies between the variation in molecular structures (encoded by
25
26 so-called molecular descriptors) and a given physico-chemical or biological property
27
28 (so-called endpoint). It is usually done for a series of often related compounds - the
29
30 dataset. These approaches have been applied successfully to predict the toxicity and
31
32 selected physico-chemical properties of a number of different types of nanoparticles
33
34 [9-15]. However, since (Q)SAR/QSPR algorithms are based on various
35
36 statistical/probabilistic approaches, sufficiently large data set must be obtained for
37
38 those methods to be functional [16]. On the other hand, engineered nanoparticles
39
40 represent very structurally diverse groups of chemicals (organic, inorganic, metal,
41
42 carbon-based nanoparticles etc.). Thus, it is difficult to build a significant dataset of
43
44 structurally related nanoparticles. In addition, since nanoparticles are not structurally
45
46 homogenous, a common mechanism of toxicity (which is a fundamental requirement
47
48 for modeling) cannot be expected for all of them. As a consequence, toxicity and other
49
50 properties should be assessed within specific applicability domains, i.e. groups of
51
52
53
54
55
56
57
58
59
60

1
2
3 sufficiently similar nanoparticles that can be expected to have similar properties.
4
5 Examples of such groups include metal oxides, various modifications of a single metal
6
7 oxide, carbon nanostructures etc. Furthermore, systematic data available in the
8
9 literature for specific groups of nanomaterials are usually of limited use in the context
10
11 of (Q)SAR/QSPR modeling and risk assessment purposes. Very often, even if the
12
13 quality of the measurements is very high, the number of data points (i.e., given
14
15 property measured for a series of differing nanoparticles) is insufficient to allow for
16
17 the development of robust models [17].
18
19
20
21
22

23
24 In the absence of relevant, reliable and sufficient data to build an appropriately
25
26 validated (Q)SAR/QSPR model, the read-across approach is an attractive and
27
28 pragmatic technique to fill data gaps [6]. The read-across has been successfully
29
30 applied to predict many properties and toxicity (i.e., teratogenicity, mutagenicity, skin
31
32 and respiratory tract sensitization effects, fish acute toxicity, etc.) of compounds, such
33
34 as carbonyl compounds, phospho-organic pesticides and polar organic compounds
35
36 [18-20]. However, so far this computational technique has not been used to fill data
37
38 gaps for nanoparticles.
39
40
41
42

43
44 The focus of this work is to provide an efficient methodology that will
45
46 eliminate shortcomings of the dataset scarcity for nanomaterials. Our approach
47
48 provides a capable tool allowing for predictions of various properties (physico-
49
50 chemical as well as toxicological) of unknown nanomaterials based on information
51
52 extracted from very few known species. It opens new opportunities to evaluate their
53
54 properties without necessity of performing expensive experimental studies on large
55
56 pool of nanomaterials.
57
58
59
60

2. Read-across - general approach

The read-across concept has received much attention in recent years since is a non-testing approach that can be used for data gap filling. In principle, read-across can be used to predict endpoint information for one, or more, chemical(s) (the so-called “target chemical(s)”) by using data from the same endpoint from another substance(s) (the “source chemical(s)”), which are considered to be ‘similar’ in some way. Similarity is usually based on structural and/or physico-chemical properties of new molecules [21]. The similarities may be based on common functional groups, common constituents or the likelihood of common precursors and/or breakdown products [21]. In other words, the known activity of one substance could be used to estimate the unknown value, for the same activity, for another substance if they have a similar chemical structure. According to OECD Guidance [21] read-across may be performed in a qualitative or quantitative manner, depending on the data used (discrete or numerical). If the presence (or absence) of a property/activity for the source chemical(s) is expressed on a discrete scale, then qualitative read-across is used to obtain a ‘yes/no’ answer for the presence (or absence) of the same property/activity for one or more target chemical(s). Whereas, quantitative read-across is used to predict the unknown value of the property/activity for the target chemical(s) based on the known value(s) of the same property/activity for source chemical(s) expressed on a numerical scale. Furthermore, as already mentioned, read-across may be performed between two chemicals or for a group of chemicals in one of four ways shown schematically in Figure 1.

1
2
3
4 [INSERT FIGURE 1]
5
6
7

8 Regardless of whether the property/activity is expressed in a quantitative or
9 qualitative manner, the procedure of read-across involves two following steps: the
10 identification of a structural feature that is common to two substances or a group of
11 substances (which are considered to be similar) and the assumption that the known
12 value of a physico-chemical property, toxicological effect or environmental fate
13 property for one chemical can be used to assess the unknown value of the same
14 property or activity for another chemical. However, as already mentioned, read-across
15 approach has not been yet applied to fill data gaps for nanoparticles.
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30

31 **3. Materials and methods**

32 The methodology applied in this study involves the following steps: (i)
33 exploration of the multidimensional space of calculated molecular descriptors in order
34 to obtain the preliminary information on structural similarities and dissimilarities on
35 the studied datasets; (ii) grouping of nanoparticles with pattern recognition techniques
36 based on structural features; (iii) performing nano-read-across analysis in a variant of
37 the many-to-many approach.
38
39
40
41
42
43
44
45
46
47
48

49 Metal oxide nanoparticles were selected as the target group for the study.
50 These compounds are of the highest priority due to the fact that metal oxide
51 nanoparticles are commonly used in nanotechnology. Cytotoxicity data (the
52 concentration of metal oxide nanoparticles that reduces bacteria viability of 50%,
53
54
55
56
57
58
59
60

1
2
3
4 EC₅₀) for 17 nano-sized metal oxides to the bacterium *Escherichia coli* were taken
5
6 from our previous study [9]. Whereas the toxicity data (the concentration of metal
7
8 oxide nanoparticles that caused a 50% reduction of the cells after 24h exposure, LC₅₀)
9
10 for 18 nano-metal oxides to a human keratinocyte (HaCaT) cell line were taken from
11
12 [15]. To make the current study directly comparable with the results obtained from the
13
14 previously described nano-QSAR models [9, 15], we utilized the same method to split
15
16 the data into training and validation sets (for more details, please refer to the
17
18 Electronic Supplementary Material, Table S3 and Table S9). The training set was later
19
20 used to identify of MeOx or a group of MeOx(s) (which are considered to be similar)
21
22 and an external validation set to evaluate the predictive ability of the nano-read-across
23
24 models. Finally, the models were applied to predict the cytotoxicity towards bacteria
25
26 *E. coli* and human keratinocyte (HaCaT) cell line respectively for untested metal oxide
27
28 nanoparticles from the selected groups.
29
30
31
32
33
34

35 The structural characteristics that quantitatively describe the variation of the
36
37 nanoparticles' structure (structural descriptors) also were taken from our previous
38
39 studies [9, 15]. Detailed lists of the descriptors calculated on the basis of small,
40
41 stoichiometric clusters, reflecting all the characteristics of fragments of the crystal
42
43 structures of particular oxides are provided in the Electronic Supplementary Material
44
45 (Table S1 and Table S7). To ensure that the influence of each variable was equivalent
46
47 (i.e. the same scale and range of all variables), the descriptors were standardized,
48
49 which means that the average value of the descriptor for the set of compounds
50
51 considered was subtracted from the descriptors and the resultant values divided by
52
53 the standard deviation, according to the formula:
54
55
56
57
58
59
60

$$z_i = \frac{x_i - \bar{x}_j}{s_j} \quad (1)$$

where:

z_i - is the transformed value of a given variable,

x_i - is the original value of a given variable,

\bar{x}_j - is the mean value of a given variable calculated across a group of all studied compounds,

s_j - is the standard deviation of a given variable calculated across a group of all studied compounds.

Two-dimensional hierarchical cluster analysis (t-HCA) was employed to search for similarities between the nanoparticles in the feature space, assuming that two NPs are similar when located close (i.e., in terms of a measure of distance) to each other [22, 23]. The similarities (and dissimilarities) of the MeOx nanoparticles studied were explored using the multidimensional space of the calculated molecular descriptors and experimentally measured cytotoxicity to bacteria *E. coli* and human keratinocyte (HaCaT) cell line, respectively. Two or more objects may form clusters indicating groups of NPs having similar property and, thus, similar biological activity/toxicity. An important advantage of applying such methods is that it enables simultaneously analyze similarities of compounds and variables, being able to identify categories (classes) of NPs. These classes may be used to predict biological activity/toxicity of other NPs that could be assigned to the same family, based on its structural similarity to other members. Two-way hierarchical cluster analysis was conducted using Euclidean distance as similarity measure and Ward's method of

1
2
3 linkage. The Euclidean distance, defined as the geometric distance in the
4
5 multidimensional space, was computed according to the formula:
6
7

$$d_{(ij)} = \sqrt{\sum_{k=1}^p (x_{(i)k} - x_{(j)k})^2} \quad (2)$$

8
9
10
11
12
13 where $x_{(i)k}$, $x_{(j)k}$ are k -coordinate values for i and j object respectively.
14
15
16

17
18 Ward's method of clustering (also known as the Ward's minimum variance) is
19
20 based on the inner squared distance of clusters, so that at each stage these two clusters
21
22 are grouped together, for which the minimum increase in the total within a group error
23
24 sums of squares is observed [24]. The final number of clusters was defined based on
25
26 Sneath's criterion of distance measure relation, i.e. $1/2D_{max}$, where D_{max} is the
27
28 maximal distance in the similarity matrix [25]. The Sneath index of cluster
29
30 significance implies that only clusters remaining compact after breaking the linkage at
31
32 $1/2D_{max}$ are considered significant and should be interpreted.
33
34
35
36
37

38
39 On the basis of these clusters, the MeOx nanoparticles from the training set
40
41 were divided into groups with similar cytotoxicity profiles. After the identification of
42
43 the chemicals considered to be analogues, the information on cytotoxicity of MeOx
44
45 available for members of each class was utilized to estimate the cytotoxicity for the
46
47 compounds from the validation set. Therefore, the data from validation set were
48
49 rescaled, using the mean and standard deviation values from the predefined similarity
50
51 classes, in order to incorporate them into the classes.
52
53
54
55
56
57
58
59
60

1
2
3
4 Finally, it was hypothesized that the cytotoxicity of metal oxide nanoparticles
5
6 to bacteria *E. coli* and human keratinocyte (HaCaT) cell line obtained from the nano-
7
8 read-across techniques should not be substantially different than the values obtained
9
10 experimentally, as well as those predicted from the nano-QSAR models. To verify
11
12 whether the hypothesis and conclusions can be extended to other properties and
13
14 groups of nanoparticles, a comparison of predictive power was performed for both
15
16 techniques. Since the method proposed in this study to fill data gaps is qualitative
17
18 rather than quantitative, the non-parametric Spearman rank correlation test [26] was
19
20 applied as a measure of the strength of the relationship between the two variables.
21
22 Spearman's rho (ρ_s), based on rank orderings, describes whether two variables are
23
24 correlated and is defined as (eq. 3):
25
26
27
28

$$\rho_s = 1 - \frac{6\sum d_i^2}{N(N^2 - 1)} \quad (3)$$

29
30
31
32
33
34 where:

35
36
37 ρ_s – is Spearman's rank correlation coefficient;

38
39
40 d^2 - is the square of the difference between ranks;

41
42
43 N - is the number of data pairs.
44

45
46 Therefore, to verify whether the predictions from the nano-read-across
47
48 technique differ significantly from the experimental values of $\log(EC_{50})^{-1}$ as well as
49
50 those predicted from the nano-QSAR modeling, the Spearman's rank correlation
51
52 coefficient was calculated as an alternative to Pearson's correlation coefficient. All
53
54 calculations within this study were carried out with the Statistics Toolbox for
55
56 MATLAB v. 7.6.0.324 [27].
57
58
59
60

4. Results

4.1. Case study 1

Since the general concept of read-across methods is based on the assumption that nanoparticles with similarities in their chemical structure will give similar toxic responses, we decided to identify those structural features that can be related to the toxicity of metal oxide nanoparticles to bacteria *E. coli*.

The starting point for estimating the mutual similarity of MeOx in terms of the toxicity of metal oxide nanoparticles to bacteria *E. coli* was a matrix of 11 structural descriptors (Table S2). Selection of independent variable(s) that define the similarity of MeOx was conducted based on the value of the Pearson's correlation coefficient calculated between the matrix of all descriptors (\mathbf{X}) and the vector of the dependent variable (\mathbf{y}) for training set. The size of the training set for the identification of structure-based categories (i.e. developing a model) was $n=10$. In contrast, the size of the test set for carrying out the external validation and estimating the error of classification for novel NPs (not previously utilized for training the model) was $k=7$. To eliminate redundancy in the structural data, we utilized for further modeling only those descriptors which had been found by us to contribute significantly to understanding the mechanism of toxicity and have a Pearson correlation coefficient with the endpoint (i.e. cytotoxicity) with an absolute value greater than 0.8. It should be mentioned, however, that sometimes a model can include several descriptors which may have a low individual correlation with activity, but in combination provide a good model. Nevertheless, we generally assume that the proposed nano-read-across

1
2
3
4 approach should be maximally simplified and based on minimal number of utilized
5
6 descriptor(s). Finally, we selected only one molecular descriptor: the enthalpy of
7
8 formation of a gaseous cation having the same oxidation state as that in the metal
9
10 oxide structure (ΔH_{Me^+}). The correlation coefficient for the ΔH_{Me^+} (describing
11
12 ionization enthalpy of the (detached) metal atoms) with the toxicity was -0.92. This
13
14 means the descriptor ΔH_{Me^+} can explain approximately 85% of the variability of the
15
16 toxicity data for the oxides ($R^2=0.85$).
17
18
19
20

21 The selected descriptor i.e., enthalpy of formation of the metal cation in the gas
22
23 phase, has been used to define groups of similar chemical species for the read-across
24
25 of *E. coli* toxicity. The Hierarchical Cluster Analysis (t-HCA) was performed using
26
27 the data matrix, where the first column corresponded to the experimentally determined
28
29 cytotoxicity of MeOx to *E. coli*, and the second represented the theoretically
30
31 calculated values of enthalpy of formation of the metal cation in the gas phase, for all
32
33 individual metal oxides from the training set. As a result, a dendrogram - a tree-like
34
35 graphic, shown graphically in Figure 2, was derived. This dendrogram displays the
36
37 linkages between the clustered objects with respect to their similarity.
38
39
40
41
42
43
44

45 [INSERT FIGURE 2]
46
47
48
49
50

51 Based on the structural similarities derived from the selected quantum-
52
53 mechanical descriptor of the compounds studied, three classes (groups) of metal nano-
54
55 oxides were identified. The classes consist of compounds with similar values of
56
57
58
59
60

1
2
3
4 cytotoxicity. The first class contained the following oxides: ZnO, CuO and Y₂O₃,
5
6 second: Bi₂O₃, In₂O₃, Al₂O₃ and Fe₂O₃, while the third group includes: SiO₂, SnO₂
7
8 and TiO₂. By analyzing the results obtained, it can be seen that compounds from class
9
10 I seem to have the highest toxicity, whereas the majority of compounds in class II and
11
12 all compounds from class III qualify as being non-toxic. The observed trends are in
13
14 accordance with the results of previous studies by other groups. For instance, Heinlaan
15
16 et al. [28] observed that the toxicity rank order for ZnO, TiO₂, and CuO to bacteria
17
18 *Vibrio fischeri* and to crustaceans *D. magna* and *Thamnocephalus platyurus* was as
19
20 follows: ZnO > CuO > TiO₂.
21
22
23
24

25
26 In the next step the predictive ability of the nano-read-across model has been
27
28 evaluated. The set of seven metal oxides (i.e., CoO, Cr₂O₃, La₂O₃, NiO, Sb₂O₃, V₂O₃,
29
30 ZrO₂) from the validation set, which have not been previously applied for
31
32 classification was used for this task. The ΔH_{Me+} values for the compounds in the
33
34 validation set were rescaled in order to incorporate them into the previously identified
35
36 classes, which contained MeOx with similar values of cytotoxicity (Table S4). Based
37
38 on information on the cytotoxicity of MeOx available for training set members of each
39
40 class, as well as on the rescaled values of the selected descriptor, the cytotoxicity of
41
42 each metal oxide in the validation set was estimated. NiO and CoO were assigned to
43
44 the first class, while the rest of the metal oxides (i.e., Cr₂O₃, La₂O₃, Sb₂O₃, V₂O₃ and
45
46 ZrO₂) were included in the second class, as determined by the limit values of
47
48 $\log(\text{EC}_{50})^{-1}$, which were respectively 2.29 and 2.82 log unit [mol/dm³] (Table S4).
49
50
51
52
53

54
55 The final stage of the evaluation of the accuracy of the data estimation was to
56
57 perform two-dimensional hierarchical cluster analysis for all 17 metal oxide
58
59
60

1
2
3 nanoparticles from the training and validation sets. The result obtained is the
4 dendrogram, which is shown graphically in Figure 3.
5
6
7

8
9
10 [INSERT FIGURE 3]
11
12

13
14
15 The dendrogram confirms that the distribution of the validation set of oxides
16 into the toxicity classes, based on the rescaled descriptor values, is highly consistent
17 with the class assignment we obtained by performing the t-HCA analysis on the
18 experimental data for the training and validation sets. It should be also highlighted that
19 based on the rescaled value of ΔH_{Me+} relevant predictions were made even for CoO -
20 the metal oxide with toxicity $(\log(\text{EC}_{50})^{-1} = 3.51)$ slightly exceeding the cytotoxicity
21 range covered by the training compounds $(3.45 > \log(\text{EC}_{50})^{-1} > 1.74)$. Thus the
22 validation stage confirmed the significance of our nano-read-across model. After
23 positive validation, the nano-read-across model was applied to estimate the
24 cytotoxicity to *E. coli* for 19 untested metal oxide nanoparticles. The predicted results
25 (Figure 4) suggest that toxicity of 19 metal oxides increases in the following order:
26
27 $\text{GeO}_2 < \text{Ga}_2\text{O}_3 < \text{Tl}_2\text{O}_3 < \text{Au}_2\text{O}_3 < \text{Yb}_2\text{O}_3 \approx \text{Mn}_2\text{O}_3 < \text{Er}_2\text{O}_3 < \text{Ho}_2\text{O}_3 < \text{Eu}_2\text{O}_3 < \text{Tb}_2\text{O}_3$
28
29 $< \text{Gd}_2\text{O}_3 < \text{Sm}_2\text{O}_3 < \text{Nd}_2\text{O}_3 < \text{Ag}_2\text{O}_3 < \text{FeO} < \text{AuO} < \text{MnO} < \text{MgO} < \text{PbO}$.
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48

49 [INSERT FIGURE 4]
50
51
52
53
54

55 In summary, the results obtained (for compounds from training and validation
56 sets) that are presented in Figure 4 indicate a very good agreement with both the
57
58
59
60

1
2
3
4 experimental data and previously predicted with nano-QSAR model results.
5

6
7 A more detailed comparison between the classes provides interesting
8 information. Figure 4 shows how the cytotoxicity of MeOx to *E. coli* varies with the
9 oxidation state of the metal for each class. Metal cations are formed by the initial loss
10 of electron(s) and many metals can form cations in several oxidation states. Since
11 cations of the metals in lower oxidation states (+2) are much less chemically stable
12 than the corresponding cation of the metal in the +4 oxidation state, thus the release of
13 metal cations having a smaller charge (n) is energetically more favorable than the
14 release of cations with larger n . This explains why the metal oxides containing a cation
15 with lower oxidation numbers (i.e., +2) exhibit the highest cytotoxicity to *E. coli*,
16 whereas the nano-sized metal oxides in the +4 oxidation state of the metal - not. The
17 above observation sheds light on the mechanism of cytotoxicity to bacteria *E. coli*.
18 However, since to the identification of classes of similar nanoparticles based on the
19 structural features used the same quantum-mechanical descriptor that has been utilized
20 to develop the final nano-QSAR model (i.e., enthalpy of formation of a gaseous cation
21 having the same oxidation state as that in the metal oxide structure), its interpretation
22 is consistent with the mechanism of cytotoxicity previously discussed by Puzyn et al.
23 [9]. As already demonstrated, the cytotoxicity of metal oxides nanoparticles to *E. coli*
24 is associated with the process of metal cations release from the particle surface and
25 decreases in the following order: $\text{Me}^{2+} > \text{Me}^{3+} > \text{Me}^{4+}$. For more details, please refer to
26 Puzyn et al. [9].
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54

55 Predictions are always less accurate than the experimental data. However, the
56 results obtained from: (1) nano-QSAR modeling and (2) nano-read-across techniques
57
58
59
60

1
2
3
4 estimation that are summarized in Figure 4, provide mechanistic interpretation of the
5 investigated toxicity and indicate an almost identical order of assignment of the
6 individual metal oxides (from validation and prediction sets) to particular classes of
7 similar toxicity profile. Moreover, we found that the Spearman's rank correlation
8 coefficient calculated between experimentally measured and estimated with nano-
9 read-across approach values of $\log(\text{EC}_{50})^{-1}$ is $\rho_S = 0.955$ (Table 1). Since the absolute
10 value of ρ_S was greater than the critical $\rho_{S(\alpha 0.05)}$, and the p-value was less than the
11 significance level of 5% ($p=0.0001$), thus the null hypothesis has been rejected and
12 may be concluded that there is a rank order relationship between the variables.
13
14
15
16
17
18
19
20
21
22
23
24
25
26

27 [INSERT TABLE 1]
28
29
30
31

32 Additionally, the relationship between the cytotoxicity of metal oxide
33 nanoparticles to *E. coli* obtained from the nano-QSAR model and estimated with
34 nano-read-across technique was found to be statistically significant. Since our
35 calculated value of $\rho_S = 0.961$ exceeds the critical value and the p-value was 0.0001,
36 we concluded that the correlation is considered to be statistically significant at the
37 95% probability level (Table 1). For more details, please refer to Electronic
38 Supplementary Material (Table S5 and Table S6). The results obtained support the
39 research hypothesis that there is a rank order relationship between the predictions from
40 nano-read-across technique and experimentally measured (or predicted from the nano-
41 QSAR model) values of $\log(\text{EC}_{50})^{-1}$.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

4.2. Case study 2

The goal of this study was to identify similar groups of metal oxide nanoparticles, to find relationships between the feature/property of NPs, to detect factors responsible for their cytotoxicity to the human keratinocyte cell line as well as to reveal discriminating parameters which determine the classification of the MeOx in different groups of similarity (or dissimilarity). To explore the data set studied and to examine the similarities of the MeOx nanoparticles, the hierarchical clustering method was used. The results presented below are based on the Euclidean distance as the similarity measure, Ward's linkage algorithm, z-transformation of data and check of the cluster significance by the Sneath index. The data set studied presents 26 of structural descriptors (including 15 descriptors derived from quantum-mechanical calculations and 11 descriptors derived from TEM image analysis) and toxicity data to a human keratinocyte (HaCaT) cell line in term of LC₅₀ for 18 different metal oxides nanoparticles. The calculated/measured parameters are listed in Table S8 and Table S9.

To define structure-based categories of MeOx the Mulliken's electronegativity of the cluster (χ^C) has been selected by employing Pearson correlation coefficient. The correlation coefficient for the χ^C with the toxicity was 0.81. This means χ^C can explain approximately 66% of the variability of the toxicity the metal oxide nanoparticles.

The dendrogram formed for the selected variable (Figure 5) reveals three main classes of metal nano-oxides that consist of compounds with similar values of cytotoxicity. The first class includes: ZnO, CoO and In₂O₃, second class contains: Bi₂O₃, Mn₂O₃ and Sb₂O₃, whereas the third class is: ZrO₂, TiO₂, SiO₂, V₂O₃. By

1
2
3
4 analyzing the results obtained, we found that the cytotoxicity of MeOx systematically
5
6 decreases when moving from class I to class III.
7
8
9

10 [INSERT FIGURE 5]
11
12
13
14

15 In this way, we obtained a classification model capable of assigning novel
16
17 compounds to an appropriate class, based on the values of selected descriptor. Since
18
19 according to OECD recommendations mechanistic interpretation is vital for
20
21 validation, in the next step the predictive ability of the developed nano-read-across
22
23 model has been evaluated with the set of eight metal oxides (i.e., Al₂O₃, Cr₂O₃, Fe₂O₃,
24
25 La₂O₃, NiO, SnO₂, WO₃, Y₂O₃) from the validation set. Based on information on the
26
27 cytotoxicity of MeOx available for training set members of each class, as well as on
28
29 the rescaled values of the selected descriptor, we have estimated cytotoxicity of each
30
31 metal oxide in the validation set. We found that La₂O₃ and WO₃ have been assigned to
32
33 the first class, Cr₂O₃, NiO and SnO₂ to the second class, while the rest of the metal
34
35 oxides (i.e., Y₂O₃, Al₂O₃ and Fe₂O) were included in the third class. We obtained
36
37 relevant estimation of the majority of MeOx that was confirmed by performing two-
38
39 dimensional hierarchical cluster analysis for all metal oxide nanoparticles from both:
40
41 training and validation sets (Figure 6).
42
43
44
45
46
47
48
49
50

51 [INSERT FIGURE 6]
52
53
54
55
56

57 Interestingly, there were false negative predictions for three oxides. The nano-
58
59
60

1
2
3
4 read-across model failed to predict the high toxic effect to human keratinocytes cell
5
6 line for the two oxides, namely: SnO_2 and Mn_2O_3 . Additionally V_2O_3 was wrongly
7
8 assigned to the non-toxic class instead of to the weak-toxic class. One possible
9
10 explanation is that all of these oxides are extremely close to the lower limit of the
11
12 appropriate toxicity class (i. e. weak-toxic and toxic class respectively for (i) V_2O_3 and
13
14 (ii) MnO_2 and SnO_2). Thus this might be a reason why they were not properly
15
16 assigned to the adequate class of toxicity. Therefore in consequence MnO_2 and SnO_2
17
18 were introduced to the weak-toxicity class instead of toxic class whereas V_2O_3 was
19
20 assigned to non-toxic class in spite it being classified previously as a weak-toxic
21
22 MeOx based on experimentally measured $\log(\text{LC}_{50})^{-1}$ values.
23
24
25
26
27

28
29 Finally, we applied the nano-read-across model to estimate the toxicity to the
30
31 human keratinocyte cell line for seven - so far experimentally untested - metal oxide
32
33 nanoparticles. Based only on structural similarity (i.e. rescaled values of selected
34
35 descriptor), we assigned FeO , PbO , PbO_2 and Gd_2O_3 to the first class determined by
36
37 the limit values of $\log(\text{LC}_{50})^{-1}$, higher than 2.51 log unit [mol/dm^3] (Table S10). CuO
38
39 was introduced to the second class determined by the limit values of $\log(\text{LC}_{50})^{-1}=2.21$
40
41 - 2.50. Whereas, MnO and MoO_3 were included to the third class with the upper
42
43 experimentally measured limit value of $\log(\text{LC}_{50})^{-1}$ equal 2.21 log unit [mol/dm^3]
44
45 (Table S10).
46
47
48
49

50
51 One more important note, the nano-read-across model obtained utilizes one of
52
53 the two descriptors (i.e. Mulliken's electronegativity of the cluster) that previously
54
55 were used for nano-QSAR model development [15]. As we have already demonstrated
56
57 in [15] electronegativity (χ), corresponding to the Fermi level, falls within the middle
58
59
60

of the forbidden gap region and mainly depends on the formal charge of the cation and ionic radius (eqs. 4-6):

$$\chi(P. u.) \approx 0.274z - 0.15zr - 0.01r + 1 + \alpha \quad (4)$$

$$\chi_{cation}(eV) \approx \frac{(\chi_{cation}(P. u.) + 0.206)}{0.336} \quad (5)$$

$$\chi_{cation}(eV) \approx 0.45\chi_{cation}(eV) + 3.36 \quad (6)$$

where: z is the formal charge of the cation, r is the Shannon ionic radius and α is a correcting term specific for each cation.

Conversely, following Portier et al. [29] it should be highlighted that, even if the formal charge is large, but it is distributed over a sizeable cationic volume, no one should expect a high value of the cation electronegativity. Thus, in the context of the Haber-Weiss-Fenton cycle, mechanistic interpretation of the developed nano-read-across model was intuitive: the increase of the cation electronegativity should result in the increase of catalytic properties of metal cations and, consequently, increase the toxicity of the metal oxide nanoparticle. For more details please refer to Gajewicz et al. [15].

By employing the Spearman rank correlation test we confirmed that the values obtained from the nano-read-across technique did not differ significantly from those measured experimentally ($\rho_S = 0.732$), as well as those predicted from the nano-QSAR model ($\rho_S = 0.866$). Since the absolute values of ρ_S were greater than the critical $\rho_{S(\alpha_{0.05})}$, and the p-values were less than the significance level of 5%, we concluded that

1
2
3
4 the present method is sufficiently accurate to fill data gaps (Table 1). For more details,
5
6 please refer to Electronic Supplementary Material (Table S11 and Table S12).
7

8
9 It should be also highlighted that, from an economical point of view, both
10
11 computational methods (i.e. nano-QSAR and nano-read-across techniques) are
12
13 acceptable, since they require a relatively small number of experimental data. In fact,
14
15 both are based on data from the literature, thus performing of any extensive empirical
16
17 work was unnecessary. However, the use of the nano-read-across approach seems to
18
19 be much more profitable, as it enables the prediction for those groups of NPs, for
20
21 which the number of experimental data is insufficient to develop appropriate nano-
22
23 QSAR model(s). For example, if the experimentally determined data for biological
24
25 activity are available only for 8-10 NPs, there are too few to calibrate and validate a
26
27 nano-QSAR model. However, it is sufficient to estimate the unknown value of
28
29 endpoint information for one, or more, chemical(s) by employing the nano-read-across
30
31 approach. This makes the proposed method a very useful computational tool,
32
33 especially when the number of data points to develop and validate nano-QSAR model
34
35 is inadequate. We hope that the proposed method will soon be verified by
36
37 experimental and additional theoretical studies.
38
39
40
41
42
43
44
45

46 **5. Conclusions**

47
48 In conclusion, it can be stated that although two-way hierarchical clustering is
49
50 usually applied at the first stage of data exploration, it can lead to many valuable
51
52 observations and conclusions. Applied to the data reflecting properties and toxicity of
53
54 nanoparticles, it allows for the classification of MeOx NPs according to their potential
55
56
57
58
59
60

1
2
3 toxic effects as well as for the identification of factors responsible for their toxicity.
4
5
6 The proposed method is the first attempt to use a two-dimensional hierarchical cluster
7
8 analysis to identify classes of similar nanoparticles based on the structural features and
9
10 then use them to estimate the biological activity for empirically untested metal oxide
11
12 nanomaterials. Based on structural similarity, we have estimated the cytotoxicity for -
13
14 so far untested experimentally - metal oxides (for which only structural descriptors
15
16 have been calculated).
17
18
19

20 The method proposed, although it does not provide quantitative information on
21
22 the cytotoxicity of nanoparticles towards *E. coli* or to human keratinocytes cell line,
23
24 nevertheless it can be used as an efficient tool for the preliminary hazard assessment
25
26 of nanomaterials, as well as to identify nanomaterials that may pose a potential
27
28 negative impact to the human health and the environment. One great advantage of
29
30 grouping and nano-read-across technique is the fact that it does not require a large
31
32 number of data to identify groups of similar compounds. Thereby, in the light of
33
34 REACH regulations the proposed method has a significant practical aspect and it may
35
36 be an alternative to the extremely time-consuming, costly and questionable from
37
38 ethical point of view animal experiments.
39
40
41
42
43

44 **Electronic Supplementary Material**

45 Electronic Supplementary Material is available in the online version, at
46
47 <http://www.springer.com> or from the author.
48

49 **Acknowledgements**

50 This material is based on research sponsored by the Polish National Science Center
51
52 (grant no. UMO-2011/01/M/NZ7/01445). The authors also thank the National Science
53
54 Foundation for support from the NSF CREST Interdisciplinary Nanotoxicity Center –
55
56 grant - #HRD-0833178. A.G. thanks the European Social Fund, the State Budget and
57
58 the Pomorskie Voivodeship Budget according to the Operational Programme Human
59
60

Capital, Priority VIII, Action 8.2, Under-action 8.2.2: 'Regional Innovation Strategy' for granting her with a fellowship in frame of the project "InnoDoktorant – Scholarships for PhD students, IVth edition".

References

1. Winkler D A, Mombelli E, Pietroiusti A, Tran L, Worth A, Fadeel B and McCall M J 2013 Applying quantitative structure-activity relationship approaches to nanotoxicology: current status and future potential. *Toxicology* **313** 15-23
2. Gordon T, Chen L C, Fine J M, Schlesinger R B, Su W Y, Kimmel T A and Amdur M O 1992 Pulmonary effects of inhaled zinc oxide in human subjects, guinea pigs, rats, and rabbits. *Am. Ind. Hyg. Assoc. J.* **53** 503-9
3. Rehn B, Seiler F, Rehn S, Bruch J and Maier M 2003 Investigations on the inflammatory and genotoxic lung effects of two types of titanium dioxide: untreated and surface treated. *Toxicol. Appl. Pharmacol.* **189** 84-95
4. Chen Y, Chen J, Dong J and Jin Y 2004 Comparing study of the effect of nanosized silicon dioxide and microsized silicon dioxide on fibrogenesis in rats. *Toxicol. Ind. Health.* **20**, 21-7
5. Savolainen K, Backman U, Brouwer D, Fadeel B, Fernandes T, Kuhlbusch T, Landsiedel R, Lynch I and Pylkkänen L 2013 Nanosafety in Europe 2015-2025: Towards Safe and Sustainable Nanomaterials and Nanotechnology Innovations. Finnish Institute of Occupational Health: http://www.ttl.fi/en/publications/Electronic_publications/Nanosafety_in_europe_2015-2025/Documents/nanosafety_2015-2025.pdf
6. Regulation (EC) No 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) No 793/93 and Commission Regulation (EC) No 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. Official Journal of the European Union, L 396/1 of 30.12.2006. Office for Official Publications of the European Communities (OPOCE)
7. van der Jagt K, Munn S, Tørsløv J and de Bruijn J 2004 Alternative Approaches can Reduce the Use of Test Animals Under REACH. *JRC Report EUR 21405*
8. Directive 2010/63/EU of the European Parliament and of the Council on the protection of animals used for scientific purposes. Official Journal of the European Union L276/33.
9. Puzyn T, Rasulev B, Gajewicz A, Hu X, Dasari T P, Michalkova A, Hwang H M, Toropov A, Leszczynska D and Leszczynski J 2011 Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nat. Nanotechnol.* **6** 175-8
10. Fourches D, Pu D, Tassa C, Weissleder R, Shaw S Y, Mumper R J and Tropsha A 2010 Quantitative Nanostructure - Activity Relationship Modeling. *ACS Nano* **4** 5703-12
11. Fourches D, Pu D and Tropsha A 2011 Exploring quantitative nanostructure-activity relationships (QNAR) modelling as a tool for predicting biological effects of manufactured nanoparticles. *Comb. Chem. High Throughput Screen.* **14** 217-25
12. Liu R, Rallo R, George S, Ji Z, Nair S, Nel A E and Cohen Y 2011 Classification NanoSAR development for cytotoxicity of metal oxide nanoparticles. *Small* **7**, 1118-26
13. Toropov A A, Toropova A P, Benfenati E, Gini G, Puzyn T, Leszczynska D and Leszczynski J 2012 Novel application of the CORAL software to model cytotoxicity of

- 1
2
3
4 metal oxide nanoparticles to bacteria Escherichia coli. *Chemosphere* **89** 1098-102
- 5 **14.** Zhang H, Ji Z, Xia T, Meng H, Low-Kam C, Liu R, Pokhrel S, Lin S, Wang X, Liao Y
6 P, Wang M, Li L, Rallo R, Damoiseaux R, Telesca D, Mädler L, Cohen Y, Zink Y I and
7 Nel A E 2012 Use of Metal Oxide Nanoparticle Band Gap To Develop a Predictive
8 Paradigm for Oxidative Stress and Acute Pulmonary Inflammation. *ACS Nano* **6** 4349-
9 68
- 10 **15.** Gajewicz A, Schaeublin N, Rasulev B, Maurer E I, Hussain S, Puzyn T and
11 Leszczynski J 2014 Towards Understanding Mechanisms Governing Cytotoxicity of
12 Metal Oxides Nanoparticles: Hints from Nano-QSAR Studies. *Nanotoxicology*
13 DOI:10.3109/17435390.2014.930195
- 14 **16.** Schultz T W, Cronin M T D, Walker J D and Aptula A O 2003 Quantitative structure-
15 activity relationships (QSARs) in toxicology: a historical perspective. *J. Mol. Struc-*
16 *Theochem.* **622** 1-22
- 17 **17.** Lubinski L, Urbaszek P, Gajewicz A, Cronin M T D, Enoch S J, Madden J C,
18 Leszczynska D, Leszczynski J and Puzyn T 2013 Evaluation criteria for the quality of
19 published experimental data on nanomaterials and their usefulness for QSAR
20 modelling. *SAR QSAR Environ. Res.* **24** 995-1008
- 21 **18.** Enoch S J, Cronin M T D, Schultz T W and Madden J C 2008 Quantitative and
22 mechanistic read across for predicting the skin sensitization potential of alkenes acting
23 via Michael addition. *Chem. Res. Toxicol.* **21** 513-20
- 24 **19.** Enoch S J, Hewitt M, Cronin M T D, Azam S and Madden J C 2008 Classification of
25 chemicals according to mechanism of aquatic toxicity: an evaluation of the
26 implementation of the Verhaar scheme in Toxtree. *Chemosphere* **73** 243-8
- 27 **20.** Enoch S J, Roberts D W and Cronin M T D 2009 Formation of structural categories to
28 allow for read-across for teratogenicity. *QSAR Comb. Sci.* **28** 696-708
- 29 **21.** Organisation of Economic Cooperation and Development 2007 Guidance of grouping of
30 chemicals, ENV/JM/MONO(2007)28
- 31 **22.** Massart D L and Kaufman L 2008 *The interpretation of analytical data by use of cluster*
32 *analysis*, Wiley, New York, USA
- 33 **23.** Vogt W, Nagel D and Sator H 1987 *Cluster analysis in clinical chemistry: A Model*,
34 Wiley, New York, USA
- 35 **24.** Ward Jr. J H 1963 Hierarchical Grouping to Optimize an Objective Function. *J. Am.*
36 *Satat. Assoc.* **58** 236-44
- 37 **25.** Sneath P H A and Sokal R R 1973 *Numerical taxonomy. The principles and practice of*
38 *numerical classification*, Freeman, San Francisco, USA
- 39 **26.** Spearman C E 1910 Correlation calculated from faulty data. *Brit. J. Psychol.* **3** 271-95
- 40 **27.** MATLAB 7.6 R2008, MATLAB® The Language of Technical Computing v. R2008.
41 The MathWorks Inc. <http://www.mathworks.com>
- 42 **28.** Heinlaan M, Ivask A, Blinova I, Dubourguier H C and Kahru A 2008 Toxicity of
43 nanosized and bulk ZnO, CuO and TiO₂ to bacteria *Vibrio fischeri* and crustaceans
44 *Daphnia magna* and *Thamnocephalus platyurus* *Chemosphere* **71** 1308-16
- 45 **29.** Portier J, Campet G, Poquet A, Marcel C and Subramanian M A 2001 Degenerate
46 semiconductors in the light of electronegativity and chemical hardness. *Int. J. Inorg.*
47 *Mater.* **3** 1039-43
- 48
49
50
51
52
53
54
55
56
57
58
59
60

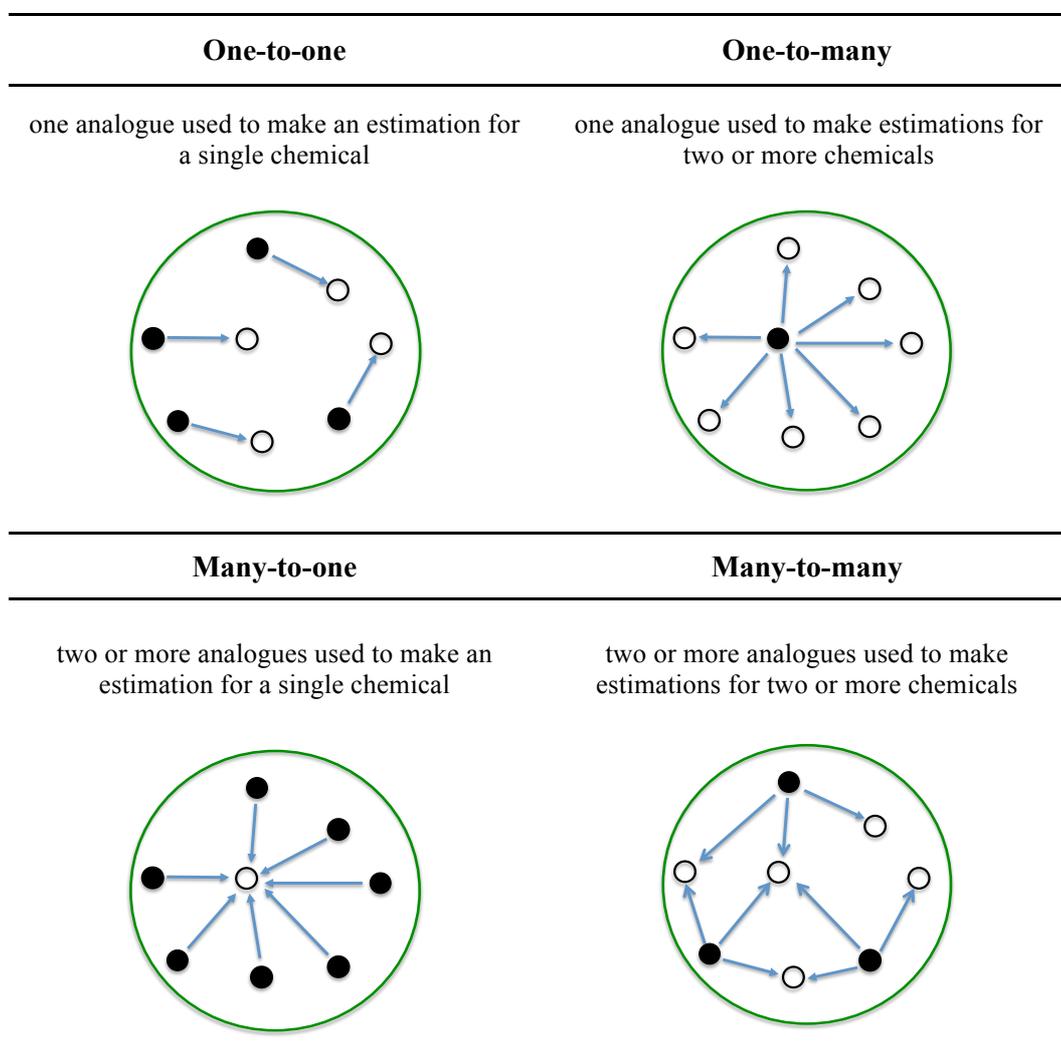
Table 1. Results of the Spearman's rank correlation test between results obtained from nano-read-across technique and experiment/nano-QSAR respectively.

Statistics	nano-read-across vs.	
	Experiment	nano-QSAR
Case study 1		
ρ_S	0.955	0.961
critical ρ_S (α 0.05)	0.414	0.279
p	0.0001	0.0001
Case study 2		
ρ_S	0.732	0.866
critical ρ_S (α 0.05)	0.401	0.337
p	0.0026	0.0001

Figure Captions

- 1
2
3
4
5
6 Fig. 1. Idea of read-across approach.
7
8
9 Fig. 2. Two-dimensional cluster analysis. Three identified natural clusters (classes)
10 in the data are presented. Colors represent the auto-scaled values of the
11 selected descriptor and cytotoxicity to bacteria *E. coli*: green color, the mean
12 value of a given descriptor/endpoint (zero value on the scale); yellow and red
13 colors, the values higher than the mean value of a given descriptor/endpoint
14 (higher up to 1.5 standard deviations); light and dark blue colors, the values
15 lower than the mean values of a given descriptor/endpoint (lower up to -1.5
16 standard deviations). When moving down the figure from Class I to Class III,
17 $\log(\text{EC}_{50})^{-1}$ systematically decreases, and (ΔH_{Me^+}) increases.
18
19
20 Fig. 3. Two-dimensional cluster analysis for all 17 metal oxide nanoparticles from
21 both: training and validation sets.
22
23
24 Fig. 4. Comparison of the predictive power of nano-QSAR and nano-read-across
25 approaches with experimentally measured values of $\log(\text{EC}_{50})^{-1}$.
26
27
28 Fig. 5. Two-dimensional cluster analysis. Three natural clusters (classes) in the data
29 were identified. Colors represent the auto-scaled values of the selected
30 descriptor and cytotoxicity to human keratinocyte cell line: green color, the
31 mean value of a given descriptor/endpoint (zero value on the scale); yellow
32 and red colors, the values higher than the mean value of a given
33 descriptor/endpoint (higher up to 1.5 standard deviations); light and dark blue
34 colors, the values lower than the mean values of a given descriptor/endpoint
35 (lower up to -1.5 standard deviations). When moving down the figure from
36 Class I to Class III, $\log(\text{LC}_{50})^{-1}$ systematically decreases, and (χ^c) increases.
37
38
39
40 Fig. 6. Two-dimensional cluster analysis for all 18 metal oxide nanoparticles from
41 both: training and validation sets.
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 1.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 2.

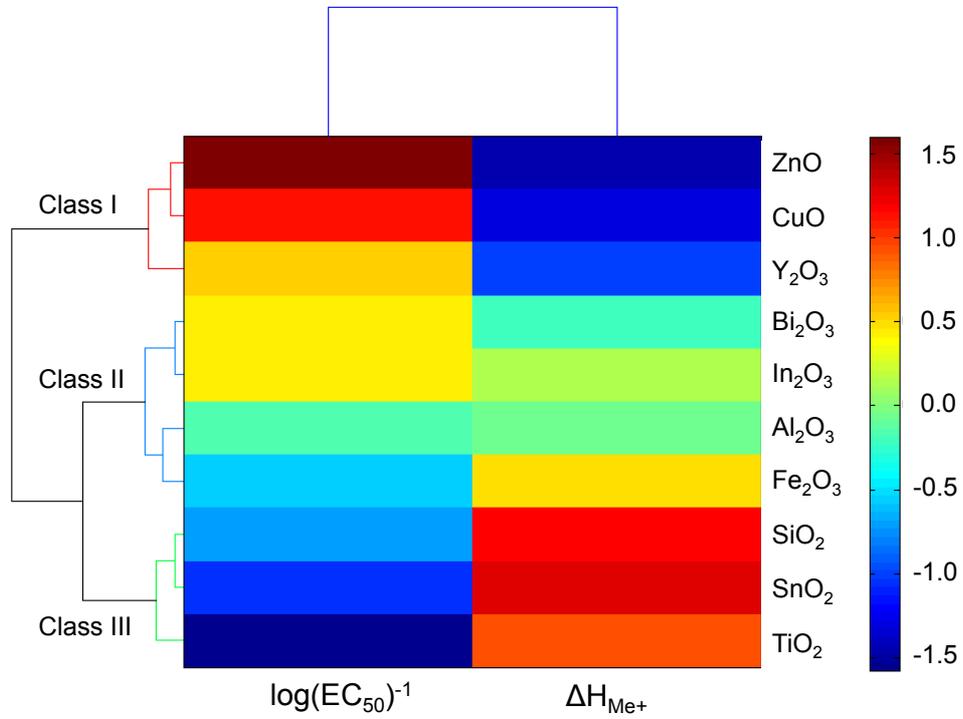


Figure 3.

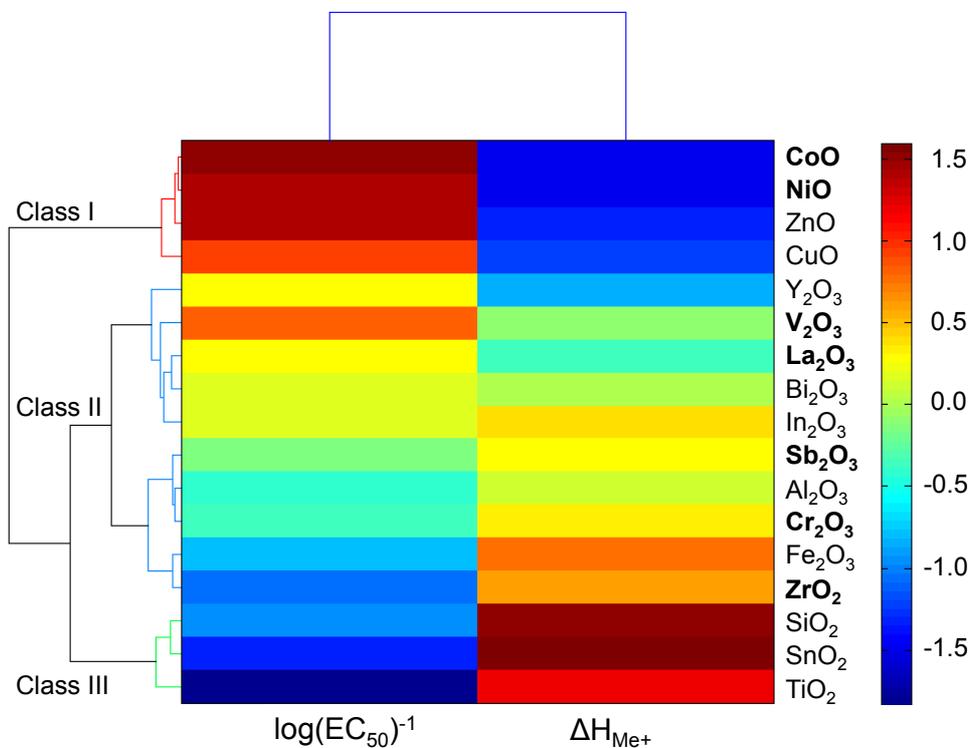
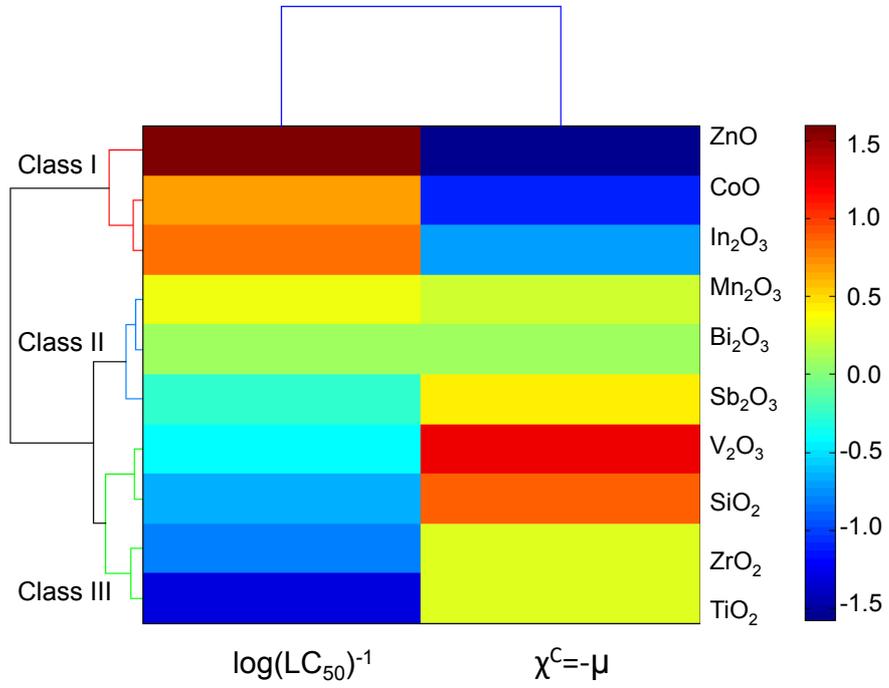


Figure 4.

Calibration and validation				Class	Prediction			
MeOx	Experimental $\log(\text{EC}_{50})^{-1}$	Predicted $\log(\text{EC}_{50})^{-1}$			MeOx	Experimental $\log(\text{EC}_{50})^{-1}$	Predicted $\log(\text{EC}_{50})^{-1}$	
		nano-QSAR	nano-read-across	nano-QSAR			nano-read-across	
CoO	3.51	3.38	CoO	Class I	PbO	N/A	3.51	PbO
ZnO	3.45	3.30	NiO		MgO	N/A	3.45	MgO
NiO	3.45	3.38	ZnO		MnO	N/A	3.44	MnO
CuO	3.20	3.24	CuO		AuO	N/A	3.23	AuO
					FeO	N/A	3.19	FeO
V ₂ O ₃	3.14	2.74	Y ₂ O ₃	Class II	Ag ₂ O ₃	N/A	3.08	Ag ₂ O ₃
Y ₂ O ₃	2.87	3.08	V ₂ O ₃		Nd ₂ O ₃	N/A	2.91	Nd ₂ O ₃
La ₂ O ₃	2.87	2.85	La ₂ O ₃		Sm ₂ O ₃	N/A	2.90	Sm ₂ O ₃
Bi ₂ O ₃	2.82	2.69	Bi ₂ O ₃		Gd ₂ O ₃	N/A	2.88	Gd ₂ O ₃
In ₂ O ₃	2.81	2.52	In ₂ O ₃		Tb ₂ O ₃	N/A	2.87	Tb ₂ O ₃
Sb ₂ O ₃	2.64	2.57	Sb ₂ O ₃		Eu ₂ O ₃	N/A	2.86	Eu ₂ O ₃
Cr ₂ O ₃	2.51	2.52	Al ₂ O ₃		Er ₂ O ₃	N/A	2.85	Er ₂ O ₃
Al ₂ O ₃	2.49	2.63	Cr ₂ O ₃		Ho ₂ O ₃	N/A	2.85	Ho ₂ O ₃
Fe ₂ O ₃	2.29	2.35	Fe ₂ O ₃		Mn ₂ O ₃	N/A	2.84	Yb ₂ O ₃
ZrO ₂	2.15	2.41	ZrO ₂		Yb ₂ O ₃	N/A	2.82	Mn ₂ O ₃
					Au ₂ O ₃	N/A	2.48	Au ₂ O ₃
					Tl ₂ O ₃	N/A	2.43	Tl ₂ O ₃
					Fe ₂ O ₃	N/A	2.40	Fe ₂ O ₃
				Ga ₂ O ₃	N/A	2.38	Ga ₂ O ₃	
SiO ₂	2.20	1.99	SiO ₂	Class III	GeO ₂	N/A	1.77	GeO ₂
SnO ₂	2.01	1.95	SnO ₂					
TiO ₂	1.74	2.13	TiO ₂					

N/A – Experimentally measured value of $\log(\text{EC}_{50})^{-1}$ is not available.

Figure 5.



1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 6.

