Reply to the comment on “Causation or only correlation? Application of causal inference graphs for evaluating causality in nano-QSAR models” by D. A. Tasi, J. Csontos, B. Nagy, Z. Kónya and G. Tasi, Nanoscale, 2018, 10, C8NR02377H

Natalia Sizochenko, a,b Agnieszka Gajewicz, a Jerzy Leszczynski b and Tomasz Puzyn a,*

The “comment” by Tasi et al.1 is related to the nano-QSAR model for predicting toxicity against E. coli published seven years ago by our group in Nature Nanotechnology2 and recently used in one of the causality analysis – the main topic of our contribution to Nanoscale: “Causation or only correlation? Application of causal inference graphs for evaluating causality in nano-QSAR models”.3

Tasi et al. expressed three critical remarks to the previously published Nature Nanotechnology model,2 including: (i) an inappropriate way of calculating the standard enthalpies of gashouse metal cation formation $\Delta H_{Me^+}$ with the semi-empirical PM6 method (the values were too high); (ii) the use of the standard enthalpy of formation $\Delta H_{Me^+}$ instead of (better correlating) ionization energy IE and (iii) the use of semi-empirical methods for calculating descriptors instead of more accurate methods of relativistic quantum chemistry (RQC).

We appreciate all the comments though they directly do not concern the ideas of the “Application of causal inference graphs for evaluating causality in nano-QSAR models” paper.3 Certainly, although our recent paper commented by Tasi et al. does not include calculation of any type of descriptor, we would like to respond to the comment, since the analysis based on literature-taken data carried out in our study does not excuse the authors from choosing them wisely.

Let’s first comment on the results obtained by Tasi et al.1 We have compared the values of the standard enthalpies of gaseous metal cation formation $\Delta H_{Me^+}$ calculated at the level of the semi-empirical PM6 method previously presented in the Nature Nanotechnology model and those calculated by Tasi et al. with the experimental data. The residuals (differences between the calculated and experimental values of $\Delta H_{Me^+}$) are presented in Fig. 1. One can notice that the values of $\Delta H_{Me^+}$ presented in Nature Nanotechnology are much closer to the experimental values than those calculated by Tasi et al. The comparison to the experiment is always the basic criterion used for evaluating the accuracy of calculations.

We have also re-evaluated the set of data presented in the comment by Tasi et al.1 using the causal inference methodology as discussed in the commented “Causation or only correlation...” paper.3 We observed mutually-driven causal relationships between the calculated ionization energies IE and the standard enthalpy of formation of gaseous metal ions $\Delta H_{Me^+}$. It is a common situation that one descriptor can be replaced in the model with another one, having similar physical...
meaning and interpretation. In Table 1 we have put together the values of Pearson’s correlation coefficients between the toxicity to E. coli modeled in the Nature Nanotechnology paper (pMIC$_{50}$) and the values of particular descriptors.

Interestingly, the correlations of the descriptor values proposed by Tasi et al.\textsuperscript{1} with toxicity were not higher than the correlation between the descriptor used in the Nature Nanotechnology model and the toxicity predictions. In other words, more accurate calculations provided by the authors Tasi et al.\textsuperscript{2} lead to the same conclusions as presented in both the original Nature Nanotechnology paper\textsuperscript{1} and our recent contribution “Causation or only correlation...”.\textsuperscript{3}

One needs to remember that in the case of QSAR-like studies, where a series (sometimes even very large) of chemicals are considered at the same time, there is always a dilemma, whether more accurate, but time-consuming or less accurate, but faster methods of calculating descriptors should be recommended. In particular, this is the case when minor differences in the accuracy level of descriptors do not significantly affect the predicted toxicity ranks. The comparison of the result of the paper by Tasi et al.\textsuperscript{1} with our models shows that one needs to consider a bigger picture while modelling complex phenomena. Application of advanced quantum chemistry methods to very small model systems might not be the most accurate way of obtaining information about the nanomaterials since it is well known that the size and shape of such systems govern their properties. On the other hand, one should be aware that each model is only a simplification of reality and as such it should be as simple as possible.

Seven years ago, when the Nature Nanotechnology model was developed, we applied the PM6 method as the most advanced semi-empirical method available at that time in MOPAC – a software package freely-distributed for academia. The choice was supported by our previous positive experience with the PM6 method.\textsuperscript{4} It should also be highlighted that the PM6 method has been parameterized to reproduce experimental reference data, which by definition considers “all possible phenomena, such as zero-point energy, internal energy, one- and two-electron phenomena, instantaneous correlation, relativistic effects, etc.”\textsuperscript{5} In addition, the PM6 method has been parameterized for all metals, which were the subject of our previous study. Nowadays, with respect to the development of relativistic methods of quantum chemistry and relativistic density functional theory on one hand, and the increasing availability of supercomputer resources on the other hand, one might contemplating the use of more advanced and accurate techniques for calculating the descriptors for metals and metal oxide nanoparticles. One should also consider that the use of the newest PM7 Hamiltonian, as declared by the developers of MOPAC, provides much more accurate results for metals and solids than the previous PM6 one.\textsuperscript{6}

In summary, although seven years after the publication of the original Nature Nanotechnology model, more accurate methods for calculating the descriptors are available, as implied by the results of the commented paper, the application of the new methods does not lead to substantial differences in the interpretation of the structural factors of metal oxide nanoparticles that affect toxicity towards bacteria E. coli. This is in contrast to what Tasi et al. have declared in the comment’s title.

### Conflicts of interest

There are no conflicts to declare.

### References