Modelling the *in vitro* cytotoxicity of metal/metal oxide and silica nanomaterials under diverse experimental conditions

A. Cassano[1], R. L. Marchese Robinson[1], A.-N. Richarz[1], M. T.D. Cronin[1]

[1] School of Pharmacy and Biomolecular Sciences, Liverpool John Moores University, Byrom Street, Liverpool, L3 3AF, United Kingdom **A.Cassano@2014.ljmu.ac.uk**

Introduction

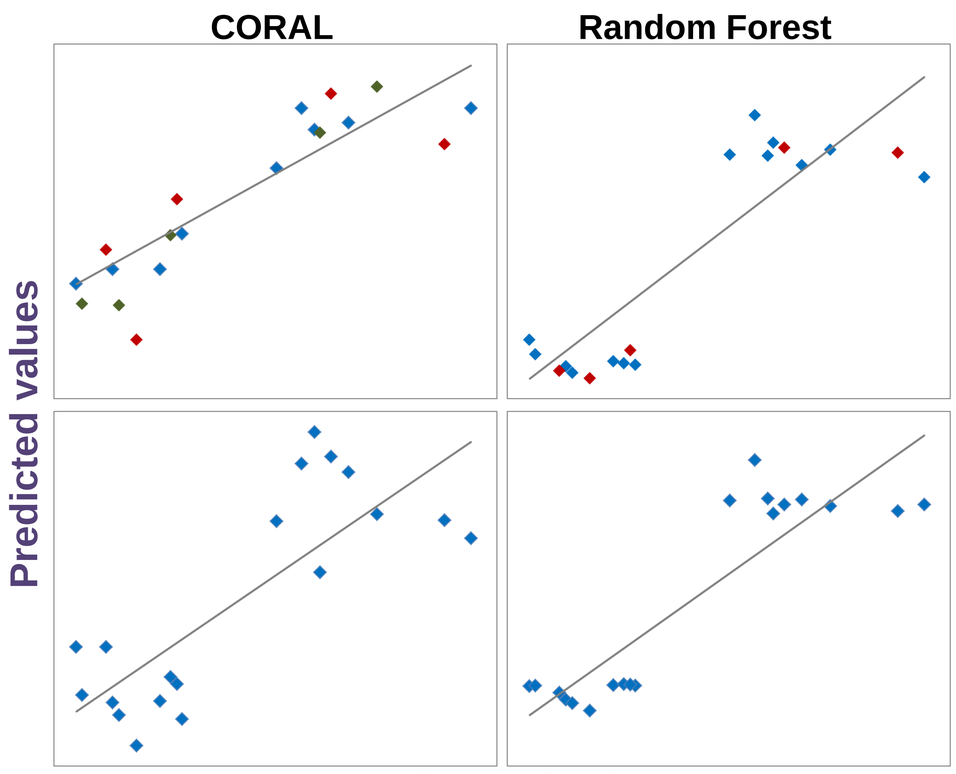
- Nanomaterials are increasingly being used in technological applications.
- However, concerns have been raised regarding their potential toxicity..
- There is considerable interest in using **computational models** to predict toxicity.
- This could enable "**safety-by-design**"
- The MODENA COST action[1] and NanoPUZZLES[2] project support the "**safety-by-design**" paradigm.
- The current work reports preliminary results of modelling **WST-1 *in vitro* cytotoxicity** experimental data for **19 uncoated silica (SiO_2)** nanomaterials kindly provided by the MODENA COST action.
- Modelling was based on the quantitative structure-activity relationship (**nano-QSAR**) approach.

The adopted nano-QSAR paradigm

- In this work, the traditional paradigm of QSAR models for chemical compounds was modified as follows:

Experimental conditions like exposure duration and cell type were used as descriptors together with descriptors derived from nanomaterials' **physicochemical properties** such as size, aspect ratio and zeta potential.

Preliminary results



Subset	CORAL R^2	Random Forest R^2
Training	0.90	0.74
Sub-training	0.86	NA
Calibration	0.98	NA
Test	0.59	0.90

Model	LOO R^2
CORAL	0.64
Random Forest	0.79

References

1. Modelling Nanomaterial Toxicity 'MODENA' COST (http://www.modena-cost.eu/)
2. NanoPUZZLES project (http://www.nanopuzzles.eu) .
3. CORAL software (http://www.insilico.eu/CORAL) .
4. A. Liaw and M. Wiener (2002). Classification and Regression by randomForest. R News 2(3), 18--22.
5. R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (https://www.R-project.org/)

Materials and methods

In the current work, the Monte Carlo method implemented in the CORAL software[3] (version: December 17, 2014) and the Random Forest algorithm implemented in the randomForest[4] R[5] package (version 4.6-12) were compared.

- **The CORAL software uses strings as input e.g. pseudo-SMILES:**

Subset	ID	Pseudo-SMILES	Exp.
+	119	AGHKM	-1.299
-	103	BDHKM	-0.872
#	123	AGHJL	0.365

CORAL (CORrelation And Logic)

- **Subset**: +/-/# stand for sub-training/calibration/test sets.
- **Pseudo-SMILES**: five descriptors derived from experimental conditions and physicochemical properties for the selected nanomaterials were coded according to the following scheme:
  i. Treatment time: 'A' = 24h; 'B' = 48h;
  ii. Cell type: 'C' = 16HBE; 'D' = A549; 'E' = HaCaT; 'F' = NRK-52E; 'G' = THP-1 macrophage;
  iii. Average size: 'H' ≤ 30 nm; 'I' > 30 nm;
  iv. Aspect ratio: 'J' = 1; 'K' > 1;
  v. Zeta potential: 'L' > -33 mV ; 'M' ≤ -33 mV;
- **Exp.**: WST-1 *in vitro* cytotoxicity: -log[EC_{25} (mm^2/ml)]

- **Bit-strings derived from the CORAL pseudo-SMILES were used as input for the R software:**

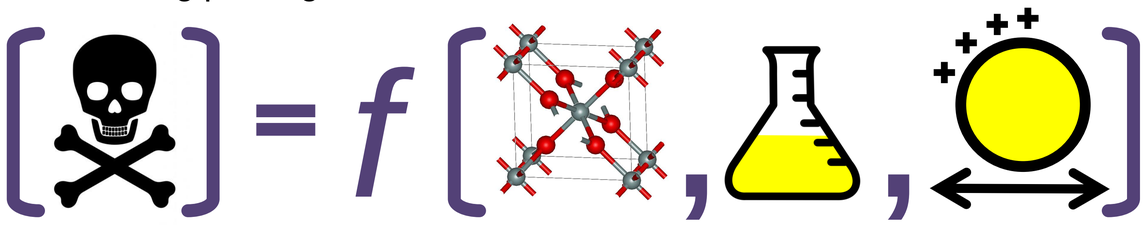
ID	Pseudo-SMILES	A	B	C	D	E	F	G	H	I	J	K	L	M	Exp.
119	AGHKM	1	0	0	0	0	0	1	1	0	0	1	0	1	-1.299
103	BDHKM	0	1	0	1	0	0	0	1	0	0	1	0	1	-0.872
123	AGHJL	1	0	0	0	0	0	1	1	0	1	0	1	0	0.365

Random Forest models were built with the default hyperparameters:
- Ntree (number of trees) = **500**
- Mtry (number of variables considered for each split) = **4**

- **Validation of modelling approaches:**
  1. Training (14 instances) and test (5 instances)
     – CORAL: training = sub-training (9) and calibration (5)
  2. "External" leave-one-out (LOO) cross-validation

Future works

- Optimise the Random Forest model hyperparameters;
- Model nanomaterials with different core/surface chemical composition (e.g. **metal oxides** and **metals**), according to the following paradigm:



- Consider additional experimental conditions such as the serum concentration, the dispersion protocol and assay type;

Acknowledgements

Funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement n° 309837 (NanoPUZZLES project) is gratefully acknowledged. The experimental data used in this work were provided by the partners of the Modelling Nanomaterial Toxicity 'MODENA' COST Initiative (COST TD1204 'MODENA'). Assistance with the use of the CORAL software was provided by Dr. Andrey Toropov and Dr. Alla Toropova.