Open Research Online

The Open University's repository of research publications
and other research outputs

Integration of Toxicity Data from Experiments and Non-Testing Methods within a Weight of Evidence Procedure

Thesis

How to cite:

Golbamaki Bakhtyari, Azadi (2018). Integration of Toxicity Data from Experiments and Non-Testing Methods within a Weight of Evidence Procedure. PhD thesis The Open University.

For guidance on citations see FAQs.

© 2018 The Author

Version: Version of Record

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data policy on reuse of materials please consult the policies page.

oro.open.ac.uk

Integration of Toxicity Data from Experiments and Non-Testing Methods within a Weight of Evidence Procedure

Thesis submitted for the Degree of Doctor of Philosophy

School of Life, Health and Chemical Sciences

Azadi Golbamaki Bakhtyari

Environmental Chemistry and Toxicology Laboratory
IRCCS - Istituto di Ricerche Farmacologiche “Mario Negri”
Milan, Italy

February 2018

Abstract

Assessment of human health and environmental risk is based on multiple sources of information, requiring the integration of the lines of evidence in order to reach a conclusion. There is an increasing need for data to fill the gaps and new methods for the data integration. From a regulatory point of view, risk assessors take advantage of all the available data by means of weight of evidence (WOE) and expert judgement approaches to develop conclusions about the risk posed by chemicals and also nanoparticles. The integration of the physico-chemical properties and toxicological effects shed light on relationships between the molecular properties and biological effects, leading us to non-testing methods. (Quantitative) structure-activity relationship ((Q)SAR) and read-across are examples of non-testing methods. In this dissertation, (i) two new structure-based carcinogenicity models, (ii) ToxDelta, a new read-across model for mutagenicity endpoint and (iii) a genotoxicity model for the metal oxide nanoparticles are introduced. Within the latter section, best professional judgement method is employed for the selection of reliable data from scientific publications to develop a data base of nanomaterials with their genotoxicity effect. We developed a decision tree model for the classification of these nanomaterials.

The (Q)SAR models used in qualitative WOE approaches mainly lack transparency resulting in risk estimates needing quantified uncertainties. Our two structure-based carcinogenicity models, provide transparent reasoning in their predictions. Additionally, ToxDelta provides better supported techniques in read-across terms based on the analysis of the differences of the molecules structures. We propose a basic qualitative WOE framework that couples the *in silico* models predictions with the inspections of the similar compounds. We demonstrate the application of this framework to two realistic case studies, and discuss how to deal with different and sometimes conflicting data obtained from various *in silico* models in qualitative WOE terms to facilitate structured and transparent development of answers to scientific questions.

To Nazanin and Letizia

Acknowledgements

This research was supported by the EU-TOXRISK project. The EU-ToxRisk project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 681002.

Disclaimer

This dissemination reflects only the author's view and the EU-ToxRisk Commission is not responsible for any use that may be made of the information it contains.

Table of Contents

List of Tables and Figures 1

Publications 4

Abbreviations 5

CHAPTER 1 7

Introduction 7

1.1 Genotoxicity, Carcinogenicity and Mutagenicity 7

1.1.1 Gene Mutation 9

1.1.2 Mutation in Cancer Cells 9

1.1.3 Base Alteration 10

1.2 Genotoxic and Carcinogenic Chemicals 11

1.3 The Animal Test(s) 12

1.3.1 Rodent Carcinogenicity Bioassay 12

1.3.2 Some Notes about Rodent Carcinogenicity Bioassay 13

1.4 Ames Test 14

1.5 Nongenotoxic Carcinogens 14

1.5.1 Modes of Action of Human Nongenotoxic Carcinogens 15

1.6 Genotoxic Carcinogens 15

1.6.1 Modes of Action of Human Genotoxic Carcinogens 16

1.7 Carcinogenic Categories of the Substances 18

1.8 Structural Alert Lists for Carcinogenicity and Mutagenicity 19

1.9 Current Hazard Identification Procedures and Related Considerations 20

1.10 (Q)SAR and REACH 21

1.11 REACH Guidelines 23

1.12 Genotoxicity Assessments in the EU Regulations 24
1.13 Read-across 25
1.14 Classification and Labelling and Chemical Safety Assessment 26
1.15 Development and Optimisation of Alternative Methods 27
1.15.1 Importance of Mode of Action and Weight of Evidence Approach 27
1.15.2 Alternative Methods for Detecting Nongenotoxic Carcinogens 27
1.16 Quantitative Structure–Activity Relationship (QSAR) 28
1.17 Software Packages for Mutagenicity and Carcinogenicity Predictions 29
1.17.1 VEGA Platform 29
1.17.2 DEREK Nexus 29
1.17.3 TOPKAT 30
1.17.4 MultiCASE 30
1.17.5 QSAR Toolbox 30
1.17.6 Toxtree 31
1.17.7 LAZAR 31
1.17.8 ACD/Tox Suite 31
1.17.9 Leadscope Model Applier 32
1.17.10 SARpy 32
1.18 Weight of Evidence 32
1.18.1 Weight of Evidence Method Classification 34
1.18.2 Weight of Evidence Application 36
1.18.3 Weight of Evidence Approach in Nanomaterials Risk Assessment 36
1.19 Abstracts of the Main Parts 39

1.19.1 Part 1- New Clues on Carcinogenicity-Related Substructures Derived From Mining Two Large Datasets of Chemical Compounds ... 39

1.19.2 Part 2 - Toxdelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances. ... 39

1.19.3 Part 3 – Genotoxicity Induced by Metal Oxide Nanoparticles: a Weight of Evidence Study and Effect of Particle Surface and Electronic Properties ... 39

CHAPTER 2 ... 41

2. Aim of the Study ... 41

2.1 Part 1 – Carcinogenicity Models ... 43

2.2 Part 2 – ToxDelta ... 45

2.3 Part 3 – Metal Oxide NMs Genotoxicity Model ... 46

CHAPTER 3 ... 48

3. Materials and Methods ... 48

3.1 Part 1 – Carcinogenicity Models ... 48

3.1.1 Carcinogenesis Data Sources ... 48

3.1.2 Comparison between the ANTARES Dataset and the ISSCAN/CGX Dataset ... 49

3.1.3 Data for Model Validation ... 49

3.1.4 Active Molecular Fragments Identification by SARpy ... 50

3.1.5 Extracting Active Fragments ... 52

3.1.6 Internal Evaluation of the Models ... 53

3.1.7 External Evaluation of the Models ... 54

3.2 Part 2 – ToxDelta ... 55

3.2.1 Database of Active and Inactive Structural Alerts ... 55

3.2.2 The MCS Algorithm ... 55

3.2.3 ToxDelta Implementation 57

3.3 Part 3 – Metal Oxide NMs Genotoxicity Model 61

3.3.1 Data Collection and Assessment 61

3.3.2 Weight of Evidence Approach in the Evaluation of the Data Set 63

3.3.3 Case Studies for Illustrating the Weight of Evidence Evaluation 64

3.3.4 Computational Analysis of nanomaterials Structure and Descriptor Generation: Quantum-chemical Descriptors 65

3.3.5 Classification SAR Modelling Methods 66

3.4 Weight of Evidence Approach in the Analysis of Results of Different In Silico Methods for the Mutagenicity Assessment of Chemicals 71

CHAPTER 4 72

4. Results and Discussions 72

4.1 Part 1- Carcinogenicity Models 72

4.1.1 R Model 72

4.1.2 E Model 72

4.1.3 Analysis of the Combination of the Prediction Results of the R and E Models 73

4.1.4 Fragments Analysis 74

4.2 Part 2- ToxDelta 79

4.2.1 Case Study 1: Benzodiazepine Derivatives 79

4.2.2 Case Study 2: Androstane Derivatives 82

4.3 Part 3 – Metal Oxide NMs Genotoxicity Model 86

4.3.1 Data Quality Assessment 86

4.3.2 Quantum Mechanical Descriptor Calculations 88

4.3.3 Recursive Partitioning and Regression Trees 89

4.4 Weight of Evidence Approach in In Silico Models for the Mutagenicity Assessment of Chemicals 92

CHAPTER 5 93

5. Conclusions 93

5.1 Part 1- Carcinogenicity Models 93

5.2 Part 2- ToxDelta 95

5.3 Part 3- Metal Oxide NMs Genotoxicity Model 97

5.4 Overall Conclusion 100

CHAPTER 6 102

6. Example of the Use of Non-Testing Methods within a Weight of Evidence Framework 102

6.1 First Case Study: Valproic Acid 103

6.2 Second Case Study: Diclofenac 111

References 121

Appendices 131

Metal Oxide nanomaterials Genotoxicity Model Supplementary Information 131

New clues on carcinogenicity-related substructures derived from mining two large datasets of chemical compounds 163

Structural Alerts for Carcinogen Compounds (R Model) 182

Structural Alerts for Carcinogen Compounds (E model) 186

ToxDelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances 189

List of Tables and Figures

Tables:

Table 1. Structural alerts belonging to certain mechanistic domains

Table 2. Test methods most commonly used for genotoxicity/mutagenicity testing

Table 3. Weight of evidence methods

Table 4. Criteria for the usefulness and quality assessment of the data set for the (Q)SAR modelling: extent of comet assay conditions checklist. General parameters have been used to assess each data point and the results are reported in Table S1.A (Appendices) where all questions are answered in a yes or no fashion.

Table 5. Comet assay experimental results for all selected metal oxide nanomaterials used for (Q)SAR modelling

Table 6. Acronyms, short definitions and units of the molecular descriptors calculated by MOPAC2012.

Table 7. R model internal and external validation for five different splits and the average of the model performance

Table 8. R model and E model internal and external validation

Table 9. The combination of the predictions of the R and E models on the ECHA external validation set

Table 10. New carcinogenic structural alerts identified by SARpy in the R and E models

Table 11. Chemicals structures in the ISSCAN-CGX data set from which structural alert 24 has been extracted

Table 12. The two case studies: Case study 1) target molecule: Diazepam, source molecule: Flunitrazepam; Case study 2) target molecule: cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-, source molecule: mepitiostane, and the results of ToxDelta: maximum common substructure and dissimilar fragments.

Table 13. Assignment of the reliability of *in vitro* comet assays based on the criteria defined in Huk et al. The assignment questions were treated in a “yes” or “no” fashion. In a weight of

evidence approach, data points that presented 1st class property were used to assign the genotoxic or non-genotoxic fate to the metal oxide nanomaterials with the same chemical core composition.

Table 14. Summary of the prediction results of non-testing models for Valproic acid

Table 15. Experimental and prediction values for some examples of similar chemicals to Valproic acid (CAS number: 99-66-1) in the training set and test set of T.E.S.T.

Table 16. Optional tabular format for summarizing weight of evidence assessment of Valproic acid

Table 17. Summary of the prediction results of non-testing models for Diclofenac

Table 18. Experimental and prediction values for N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS 66789-80-8) and N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9) as examples of similar chemicals to Diclofenac in the outcome of T.E.S.T.

Table 19. The mutagenicity and non-mutagenicity rules identified by ToxRead in the structure of Diclofenac

Table 20. Optional tabular format for summarizing weight of evidence assessment of Diclofenac

Figures:

Figure 1. Multistage carcinogenesis

Figure 2. Parallels between recombination and certain types of mutational repair.

Figure 3. Pairing between the normal form of the bases

Figure 4. Alkylation-induced specific mispairings

Figure 5. (a) Structures of common intercalating agents and (b) their interaction with DNA

Figure 6. The MCS between two molecules is shown with bold lines, and the other branches are the differences

Figure 7. The flow chart of ToxDelta: the molecular similarity/dissimilarity structure analysis software for the mutagenicity endpoint

Figure 8. The 'rpart' classification tree model on the data set of 16 metal oxide nanomaterials.

Figure 9. Valproic acid chemical structure (CAS number: 99-66-1)

Figure 10. A) ToxDelta outcome for the comparison between Valproic acid (molecule #1) and 2-[Ethyl(nitroso)amino]ethanol (CAS 13147-25-6) (molecule #2) The maximum common substructure is at the top of the panel. The dissimilar substructures are listed below their corresponding molecule. B) The identified dissimilar substructure extracted from the molecule #2 is a mutagenicity structural alert in the CRS4 dataset of mutagenicity ruleset with accuracy=1.

Figure 11. ToxRead chart for the target molecule Valproic acid (CAS number: 99-66-1). The numbers refer to CAS identifiers. Straight arrows link the target chemical to rules, while curved arrows link to chemicals

Figure 12. Diclofenac, CAS number: 15307-86-5

Figure 13. ToxDelta outcome for the comparison between Diclofenac and a) N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS 66789-80-8), b) N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9)

Figure 14. ToxDelta outcome of the dissimilar substructure remaining after the subtraction of the MCS and the structural rules identified inside the dissimilar substructure of N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9) (molecule #2)

Figure 15. ToxRead chart for the target molecule Diclofenac. The target substance molecule is shown on the right and the identified non-mutagenic and mutagenic rules are listed below

Figure 16. (a) SM197, Diphenylamine: non-mutagenic rule, (b) MNM37, Phenylamine: mutagenic rule identified by ToxRead

Figure 17. The two similar chemicals to Diclofenac extracted by ToxRead with mutagenic effect

Publications

Golbamaki, A., et al. "Comparison of in silico models for prediction of Daphnia magna acute toxicity." *SAR and QSAR in Environmental Research* 25.8 (2014): 673-694.

Gini, G., et al. "ToxRead: a tool to assist in read across and its use to assess mutagenicity of chemicals." *SAR and QSAR in Environmental Research* 25.12 (2014): 999-1011.

Cappelli, C. I., et al. "Assessment of in silico models for acute aquatic toxicity towards fish under REACH regulation." *SAR and QSAR in Environmental Research* 26.12 (2015): 977-999.

Golbamaki, A., et al. "Classification nano-SAR modeling of metal oxides nanoparticles genotoxicity based on comet assay data." Toxicology Letters 258 (2016): S271.

Golbamaki, A., et al. "New clues on carcinogenicity-related substructures derived from mining two large datasets of chemical compounds." *Journal of Environmental Science and Health, Part C* 34.2 (2016): 97-113.

Golbamaki, A., and Emilio Benfenati. "In Silico Methods for Carcinogenicity Assessment." In *In Silico Methods for Predicting Drug Toxicity* (2016): 107-119.

Golbamaki, A., et al. "The Maximum Common Substructure (MCS) Search as a New Tool for SAR and QSAR". In *Advances in QSAR Modeling* (pp. 149-165). Springer, Cham.

Golbamaki, A., et al. "ToxDelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances." *Drug Des* 6.153 (2017): 2169-0138.

Abbreviations

Best Professional Judgement (BPJ)

Bio-Concentration Factor (BCF)

Carcinogenic Potency Database (CPDB)

Classification Labelling and Packaging (CLP)

Counter Propagation Artificial Neural Network (CP ANN)

Density Functional Theory (DFT)

Distributed Structure-Searchable Toxicity (DSSTox)

EU Regulation for Registration, Evaluation, Authorization and Restriction of Chemicals (REACH)

European Chemical Agency (ECHA)

European Food Safety Authority (EFSA)

European Union Reference Laboratory for Alternative to Animal Testing (EURL ECVAM)

Food and Drug Administration (FDA)

Gap Junction Intercellular Communication (GJIC)

Health and Environmental Science Institute's (HESI)

International Agency for Research on Cancer (IARC)

International Life Sciences Institute (ILSI)

Lowest Observed Adverse Effect (LOAEL)

Matthews Correlation Coefficient (MCC)

Maximum Common Substructure (MCS)

Maximum Tolerated Dose (MTD)

Nano Material (NM)

Nano Particles (NP)

National Toxicology Program (NTP)

Organization for Economic Cooperation and Development (OECD)

Parameterized Model 7 (PM7)

Pharmaceuticals for Human Use (ICH)

(Quantitative) Structure-Activity Relationships ((Q)SAR)

Reactive Oxygen Species (ROS)

Simplified Molecular Input Line Entry Specification (SMILES)

Structural Alert (SA)

Structure-Activity Relationships (SAR)

United Nations Globally Harmonized System (UN-GHS)

Veterinary Medicinal Products (VICH)

Weight of Evidence (WOE)

CHAPTER 1

Introduction

1.1 Genotoxicity, Carcinogenicity and Mutagenicity

"The term *carcinogen* denotes a chemical substance or a mixture of chemical substances which induce cancer or increase its incidence" [1].

Carcinogenicity is a crucial endpoint for the chemical safety. Carcinogenic compounds may promote carcinogenicity in one of the three phases of causing cancer: initiation, promotion and progression [2] (Figure 1-page 7). Carcinogenesis begins with a mutation, a change of a genetic material for which no DNA repair mechanism during cell proliferation has happened. This happens in the initiation phase. During the second phase (promotion) which is reversible, the initiated cells are affected by endogenous or exogenous chemicals and because of the clonal growth, the tumour starts to form. For this reason these endogenous or exogenous chemicals are called promoters. These chemicals are not intrinsically mutagenic but cause changes in gene expression or other mechanisms that will be passed to the daughter cells. At this point, cell proliferation rate increases and apoptotic cell death decreases. In the last stage (progression) additional genotoxic events such as chromosomal aberrations and translocations take place. Progression is irreversible and it leads to the formation of neoplasms, benign and malignant alike [3–5].

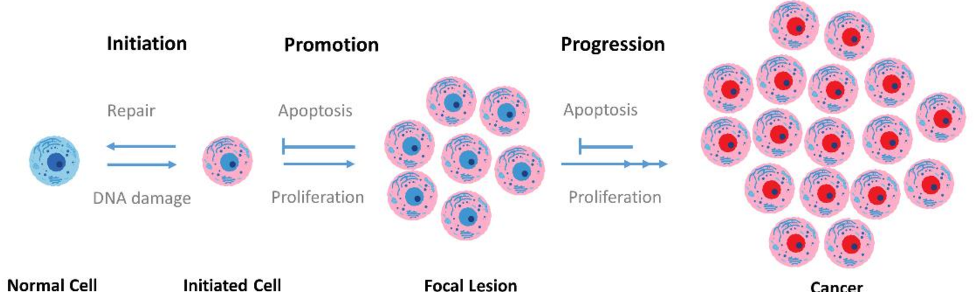


Figure 1. Multistage carcinogenesis

Genotoxicity describes a damaging action on a cell's genetic material affecting its integrity. Genotoxicity is similar to mutagenicity except that genotoxic effects that cause DNA damage are not themselves necessarily transmissible to the next generation of cells, while mutagenicity

refers to the production of transmissible genetic alterations. Genotoxic substances which are capable of causing genetic mutation (pre-mutagenic) and contributing to the development of tumours (carcinogenic) are known to be potentially mutagenic or carcinogenic. Certain chemical compounds and some radiations can induce genotoxicity.

Even low exposure levels of genotoxic substances may actuate serious health effects in somatic and germ cells. Somatic cell genotoxicity plays a role in a variety of genetic diseases. Also degenerative conditions such as accelerated aging, immune dysfunction, cardiovascular and neurodegenerative diseases and cancer are the outcome of accumulation of DNA damage in somatic cells. Mutations in germ cells can lead to spontaneous abortions, infertility or heritable damage to the offspring and possibly to the subsequent generations.

There is a strong correlation between mutagenicity and carcinogenicity. Studies show that approximately 90 percent of the known carcinogens are also mutagens. The somatic mutation theory of cancer states that the mutation of the somatic cells cause cancer.

According to the mode of action, carcinogens can be classified into genotoxic or nongenotoxic carcinogens. Genotoxic carcinogens interact directly with DNA, resulting DNA damage or chromosomal aberrations that can be detected by genotoxicity tests [6]. Adversely, nongenotoxic carcinogens have no direct reactivity with DNA and use other mechanisms in the process of tumour development such as affecting gene expression, signal transduction, and/or cell proliferation.

Mutation may occur in two modes: "spontaneously" or "inducted mutagenicity". DNA molecules are not stable in the cellular environment and each base pair in a DNA double helix is mutable with a certain probability. Mutations may affect entire chromosomes or large pieces of chromosomes. Gene alterations are the simplest form of mutation. This gene alteration is swapping of one base pair for another. Another cause of mutation can be the insertion of a transposable element from outside the genome. Most of the time the DNA damages are identified and corrected by cells. Figure 2 (page 9) shows the parallels between crossing-over and two kinds of mutational repair (excision and double-strand break repair).

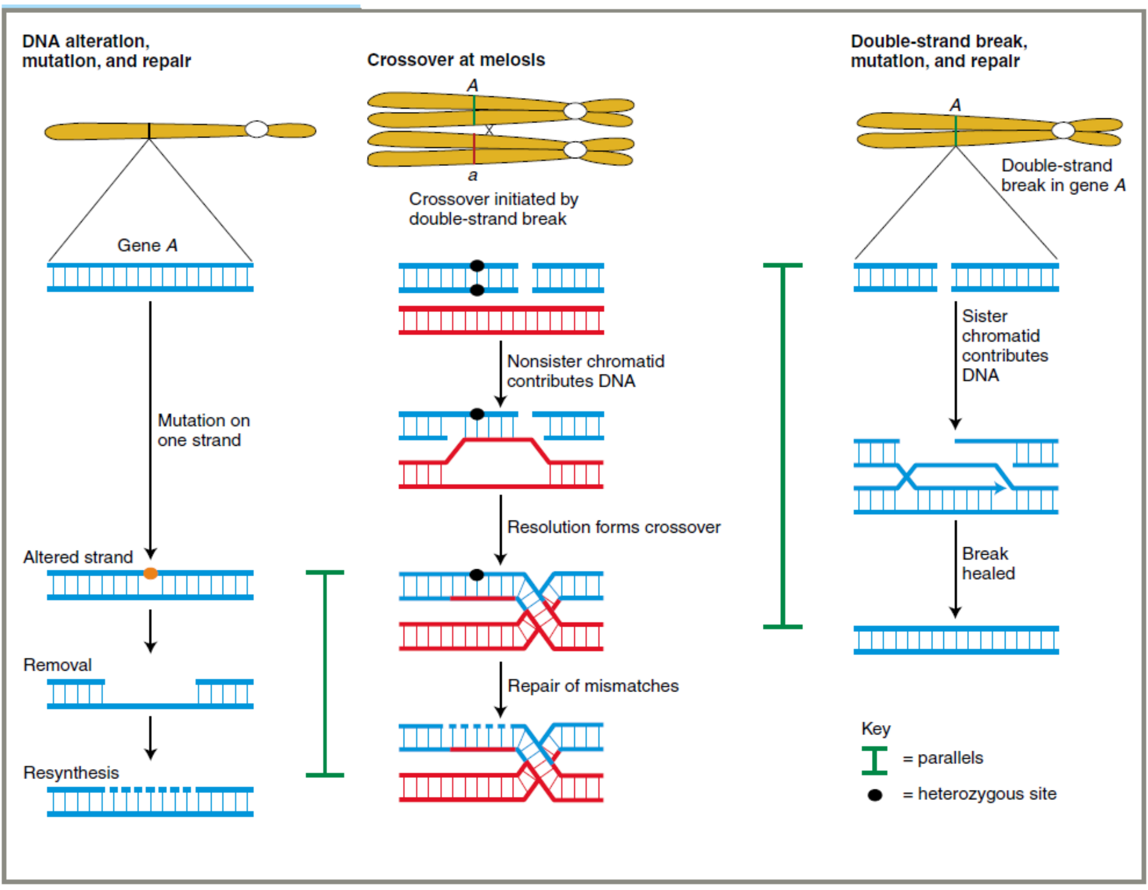


Figure 2. Parallels between recombination and certain types of mutational repair.[7]

1.1.1 Gene Mutation

The gene mutation can be divided into two classes:

- Mutations affecting single base pairs;
- Mutations altering the number of copies of a small repeated sequence within a gene.

1.1.2 Mutation in Cancer Cells

Tumours occur from a sequence of mutational incidences that lead to uncontrolled proliferation and cellular immortality. The transformation of cells from the benign into the carcinogenic state has genetic origins.

1. Most of the induced carcinogens (chemical substances and radiations) are also mutagenic and they cause cancer by originating mutations into cells.
2. A large number of mutagens affiliated with cancer have been identified. Experimental models (*in vivo* and *in vitro*) help to find these associations between mutagenicity and carcinogenicity.

1.1.3 Base Alteration

Sometimes the mutagenic agent alters a base in a DNA causing specific mispairing. Figure 3 (page 10) shows the pairing between the normal forms of the bases. Certain alkylating agents such as ethylmethanesulfonate and nitrosoguanidine, are some of the examples of these mutagens that operate by this pathway. These mutagens add alkyl groups to many positions on all four bases. Figure 4 (page 10) shows the alkylation that leads to direct mispairing and results in G.C -> A.T transitions in the next round of replication. Another important class of DNA modifiers are intercalating agents. Compounds such as proflavin, acridine orange and ICR compounds are some examples of this group (Figure 5-page 11). These agents are able to intercalate between the stacked nitrogen bases at the core of the DNA and cause single-nucleotide-pair or deletions.

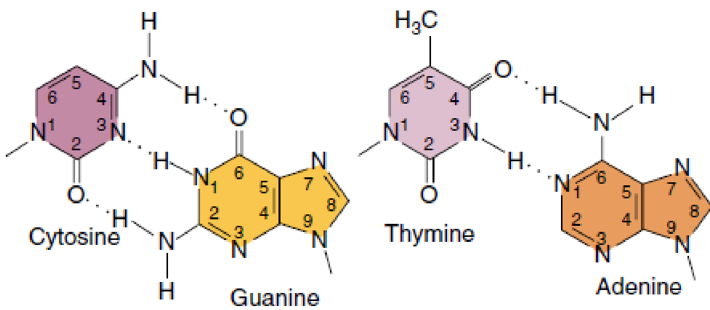
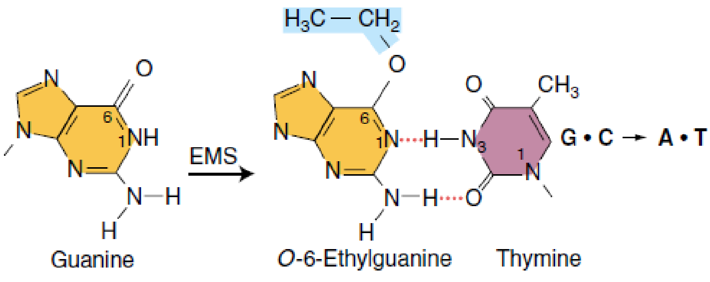


Figure 3. Pairing between the normal form of the bases [7]



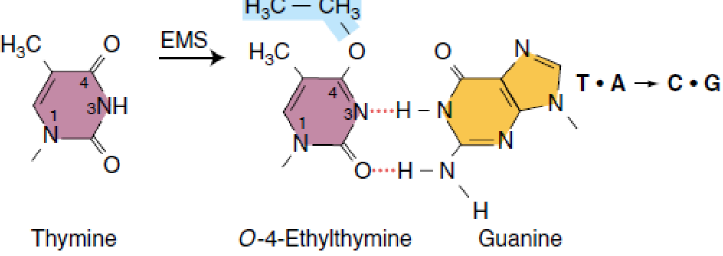


Figure 4. Alkylation-induced specific mispairings [7]

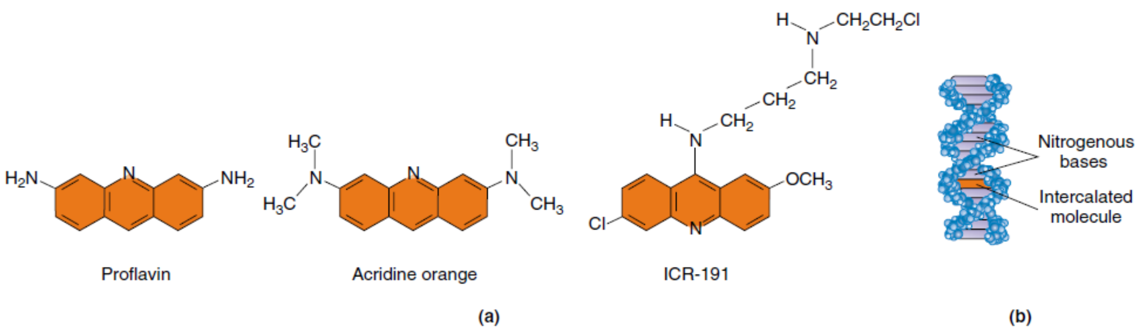


Figure 5. (a) Structures of common intercalating agents and (b) their interaction with DNA [8]

1.2 Genotoxic and Carcinogenic Chemicals

Absorption ways of chemical carcinogens following exposure are oral, inhalator, cutaneous and injection. Afterwards, different issues will be involved [9]. All the substances absorbed orally are distributed in the body through the liver, whereas those absorbed by the lung will enter the blood and after that will reach the liver [10]. Genotoxic chemical carcinogens directly damage DNA, non genotoxic carcinogens or procarcarcinogens require enzymatic conversion before affecting DNA [11].

Non genotoxic chemical carcinogens require bioactivation to electrophiles in order to bind covalently to DNA and often act by producing mutations. Different enzymes are involved in bioactivation reactions, such as oxidation, reduction, thiol conjugation, acetyl transfer, sulfur transfer, methyl transfer, glucuronosyl transfer, and epoxide hydrolysis. These enzymes are classified as oxidoreductases [12]. Human body controls metabolic activation by phase I reactions. Phase II reactions protect the body by transformation of activated compounds into inert products that will be eliminated from the body [13]. Phase II enzymes have role in the conjugation and inactivation of carcinogens and include transferases. Originally, these enzymes were considered to be involved only in the detoxification of biotransformation, but they can also trigger activation of certain epigenetic carcinogens [11].

Peroxidations occur together with metabolic reactions and cause production of ROS [14,15]. Several chronic diseases are related to these radicals including chemical carcinogenesis [14]. Chemical reactions such as oxidation, nitration/nitrosation and halogenation, which are associated with ROS trigger damage to DNA, RNA, and proteins. Consequently, mutations and alterations in the functions of important enzymes and proteins occur as a result [13]. It is

demonstrated that excess amount of ROS created by chemical compounds stimulates initiation, promotion and progression of tomour through genotoxicity [16,17].

Although the above-mentioned metabolic methods are important for both humans and animals, differences are important to be considered. Incorrect interpretations may occur in animal models used in the assays and analysis of chemicals carcinogenicity [18,19].

1.3 The Animal Test(s)

Carcinogenicity studies cannot be limited exclusively to the epidemiological data about carcinogens. For over 40 years, long-term rodent carcinogenicity bioassay using the maximum tolerated dose in 2 species over 2 yr. has been the standard procedure for detecting potential human carcinogens[20].

Data obtained from the initial 2-year carcinogenicity studies is often subjected to critiques of the screening procedures since it is inadequate for risk assessment regulatory decisions.

In laboratory experiments on animals, it is shown that most potent mutagenic chemicals are also carcinogenic[21]. Thus, all the chemicals that are mutagenic in animals are considered also mutagenic or suspected mutagenic and consequently human carcinogens, until there is found some reliable evidence which shows the contrary [22].

1.3.1 Rodent Carcinogenicity Bioassay

There is increasing understanding that carcinogenesis is a multistep process [23–25]. Chemicals with positive carcinogenicity results are subjected to more accurate and detailed evaluations about their unsafe effects for humans. Verifying whether or not the “carcinogen” chemical with positive results in long-term rodent assays is also hazardous for humans needs more chemical evaluations. These evaluations are necessary to understand the dose-response relationships, the potential hazard for humans [26]. It may be unrealistic to expect a basic 2 year study to provide all the complex data needed for risk management decisions.

Even though the quality of the studies used in toxicity assessments is high, these toxicity conclusions are not sufficient for the regulatory agencies. Some results in the NCI/NTP data base and other data sources clearly show that some chemicals that cause cancer in rodent models are not carcinogenic for humans [23,27]. The procedure which states that a chemical with

positive effect in at least one of the 4 sex/species combinations is "carcinogen", is more adequate for selecting compounds to undergo further study than regulatory purposes. Dose extrapolation, different organ responses are among the essential factors that influence the applicability of the rodent bioassays directly to risk assessment.

1.3.2 Some Notes about Rodent Carcinogenicity Bioassay

Mechanistic considerations are essential in carcinogenicity studies especially for nongenotoxic chemicals. Use of cell culture methods will fill the gap of information about the differential metabolism between animal models and humans. According to the legislations for risk assessment, data obtained from different sources, included experiments and mechanistic studies need to be used for the decision-making process. Many nongenotoxic chemicals are sex- or species specific, for this reason the mechanism of tumour formation has to be studied in both species and sex [28].

The reasonable solution is considering and using the whole available data rather than relying only on the most sensitive test results [29]. Information about the chemical concentration used for each animal species or sex combination is crucial and explains the sex/species specificity of the chemical effects [29]. This additional information plays an important role in extrapolation of the results to humans. Studying specific chemicals by rodent tests, produces useful mechanistic information. The methodologies for predicting carcinogenicity can be explored by conducting high quality rodent studies. These studies will lead us to developing better dose-response relationships and increasing our knowledge about interspecies extrapolation. Despite all the progressions in the animal tests, there is still lack of adequate rodent studies for identifying carcinogens [30].

The 2-year rodent studies are the most expensive tests that usually take place as the first step of the carcinogenicity assessment of a chemical. The following assessments, in case the rodent test result is positive, are more mechanistic, quicker and less expensive. The use of predictive models prior to the animal studies should be more reliable in the chemical evaluation process. This approach makes more resources available for the mechanistic studies and will accelerate the risk assessment deliberations for humans [31].

1.4 Ames Test

Tests for carcinogenicity are generally time consuming and are performed on small mammalian animals. Alternative tests use microbes (e.g. fungi and bacteria) and test for mutagenicity instead of carcinogenicity. Any living organism can be used for testing the mutagenicity of a chemical, this is because DNA is chemically equal in all organisms. Bacteria can be used as an alternative to mammalian models, as its life cycle is much shorter and the results can be obtained easier and faster. The most famous mutagenicity test was developed by Bruce Ames in the 1970s, which is done using Salmonella typhimurium. Properties of the bacteria were genetically engineered into these strains to make them suitable for mutagen detection. The genotype of the mutant strains in this assay is given as his-. In addition, they carry a mutation that eliminates the protective lipopolysaccharide coating of wild-type Salmonella to facilitate the entry of many different chemicals into the cell.

In a media lacking histidine this mutant bacteria will die. The “revertant” mutants revert the his – to his + genotype and phenotype and this will help the bacteria to grow in a media without histidine. In the Ames test the Salmonella bacteria is placed on plates with a very small amount of histidine and the chemical to be tested is added to the plate. The grown colonies on the plate indicate the number of revertants. To generate a dose-response curve, different concentrations of the chemical under study is tested.

A positive result of the mutagenicity Salmonella typhimurium test is an indication of the high probability that the tested chemical will be carcinogenic in laboratory animals and in consequence is more likely to be a carcinogen. Not all chemicals that cause cancer in laboratory animals are mutagenic in the Ames test, but still three quarters of the chemicals with positive result in Salmonella test are carcinogenic in also animal studies. The rapidity (3-4 weeks) and low cost of the Ames test makes it an important tool for the mutagenicity screening.

1.5 Nongenotoxic Carcinogens

A great number of human carcinogens are “genotoxic” chemicals, which means their carcinogenicity effect is caused by inducing DNA damage. The rest of the carcinogens are named “nongenotoxic” chemicals, and they induce cancer in other modes of action. Nongenotoxic mechanisms are not as extensive as for genotoxic carcinogens, but evidence

shows that alteration in multiple pathways is responsible for their carcinogenic behaviour. Some other processes in which the nongenotoxic carcinogens act are: tumor promotion, endocrine modification, immune suppression, and tissue-specific toxicity and inflammatory responses [32,33]. Nongenotoxic carcinogens unlike genotoxic carcinogens are tissue and species specific. In the past, the unique method of identifying nongenotoxic chemicals was the 2-year carcinogenicity bioassay, but the REACH legislation recommends fewer bioassays to be used in the process of carcinogenicity assessment. The main assessment strategy of REACH for the carcinogenicity endpoint is based on Ames mutagenicity test, genotoxicity in mammalian cells (*in vitro* and *in vivo*), and germ cell mutagenicity tests. These kind of tests are unable to identify the nongenotoxic carcinogens, the result of these tests are negative for such substances [34]. Thus, it is important to understand the mechanisms of action of these nongenotoxic carcinogens in order to help the decision makers in detecting these substances.

1.5.1 Modes of Action of Human Nongenotoxic Carcinogens

Nongenotoxic carcinogens induce cancer without altering DNA, by indirect stimulation of hyperplastic responses, or chromosome number or structure. The modes of action of these chemicals include receptor and non-receptor - mediated endocrine modulation, tumour promoting, inducers of tissue-specific toxicity and inflammatory responses, immunosuppressants, or gap junction intercellular communication inhibitors. The identification of these substances is very challenging. Also the kinetics of human risk assessment is different from genotoxic chemicals, and a non-linear approach (threshold) is applied for nongenotoxic carcinogens. Because of the variety of the mechanisms of action of nongenotoxic carcinogens, the assessment is done on gathered data with a WOE approach. The assessment is done individually from 90-day toxicity studies, toxicokinetic and disposition studies. If any data about 2-year chronic bioassays in rodents and human epidemiological data is available they are also used in the WOE process.

1.6 Genotoxic Carcinogens

Genotoxic carcinogens involve direct damage to DNA, to which the cell responds by repair of the damages, arrest of the cell cycle or induction of apoptosis.

1.6.1 Modes of Action of Human Genotoxic Carcinogens

1.6.1.1 Electrophilic Chemical Reaction Mechanisms Forming Adducts with DNA

Conjunction, substitution and addition are three classical chemical reactions through which the electrophiles react with biological nucleophiles. During these mechanisms of action electron-rich component interacts with the electron-deficient one [35]. Among all the known mechanisms of covalent binding, only the following mechanisms can lead to cancer: S_N1, S_N2, acylation, Schiff base formation, Michael addition, and S_NAr. These mechanisms are used for the classification of the electrophiles into appropriate mechanistic domain (Table 1-page 17) [36].

Table 1. Structural alerts belonging to certain mechanistic domains [36–38]

Mechanistic domains	Structural alerts
S_N2	Akyl esters of either phosphonic or sulphonic acids
	Monohaloalkenes
	S- or N-mustards
	Propiolactones and propiolsultones
	Epoxides and aziridines
	Aliphatic halogens
	Alkyl nitrite
S_N1	Aromatic nitro groups
	Alkyl hydrazines
	Alkyl and aryl N-nitroso groups
	Aliphatic N-nitro group
	Aromatic nitroso group
	Aromatic amines and hydroxylamine
	Halogenated polycyclic aromatic hydrocarbon (PAH)
	Halogenated dibenzodioxins
Acylation	Aromatic diazo groups
	Acyl halides
Schiff Base Formation	Simple aldehydes
	N-methylol derivates
Michael addition	Quinones
S_NAr	Aromatic N-oxides,
	Aromatic mono- and dialkylamino groups
	Halogenated benzene

1.6.1.2 Epigenetic Mechanisms of Carcinogenic Molecules

Epigenetic chemicals cause cancer without changes in the nucleotide sequences. Epigenetic factors can be found in cells under stress. The nongenotoxic (or epigenetic) carcinogens do not

make changes in DNA and do not form DNA adducts, but they changes the expression of certain genes [39]. Epigenetic factors mainly cause cancer in two ways: by methylation or post-translational modifications of histones (acetylation). DNA methylation happens in the promoter region [40] and results in the conversion of cytosin to 5-methylcytosine, with a high mutagenic potential. Acetylation of histones is controlled by histone acetyl transferases, which are important in chromatin transformation and the regulation of gene transcription [41].

1.6.1.3 Other Factors Determining the Carcinogenic Potential of Chemical Compounds

Carcinogenicity and mutagenicity are not caused only by SAs. The presence of a SA does not imply the mutagenic or carcinogenic property in a molecule. In fact, some SAs are not metabolically active in some chemicals. Molecular weight and the size of chemicals are important factors which may make the molecule lose its toxic property. Molecules with higher weights have less chance to be absorbed by cells. State of matter may make it difficult for the compound to reach the critical point. Solubility is another factor that affects the carcinogenic or mutagenic properties of the chemicals. High hydrophilicity leads to less absorption by the cells. Geometry of chemical and chemical reactivity are other important factors [35]. There are also other factors that cause the increase of decrease of carcinogenicity and mutagenicity of the chemical compounds, such as stability and transport through the membrane and half-life [2,42].

1.7 Carcinogenic Categories of the Substances

The substances classified in the Category 1A are known or presumed human carcinogens for which their mutagenicity has been proved in epidemiological and/or animal studies.

First category is known or presumed human carcinogens. A substance is classified in category 1 for carcinogenicity on the basis of epidemiological and/or animal data.

Category 1A

Substances known to have carcinogenic potential for humans. The classification in this category is largely based on human evidence, human studies that establish a causal relationship between human exposure to a substance and the development of cancer.

Category 1B

Substances presumed to have carcinogenic potential for humans. The classification in this category is largely based on animal evidence, animal experiments for which there is sufficient evidence to demonstrate animal carcinogenicity.

Second category: category 2

Suspected human carcinogens. The placing of a substance in category 2 is done on the basis of evidence obtained from human and/or animal studies, but is not sufficiently convincing to place the substance in category 1A or 1B.[34]

1.8 Structural Alert Lists for Carcinogenicity and Mutagenicity

John Ashley introduced SAs for the first time in 1985. The SAs are molecular substructures which are associated with carcinogenicity or mutagenicity properties of the molecules. These moieties represent potential mutagenicity or carcinogenicity and are the results of a long series of studies on the mechanisms of action of the mutagenic and carcinogenic chemical compounds [43]. The SAs are useful in the prediction of toxicity and the classification of potential carcinogens, as well as, in understanding the mechanism of genotoxicity [39,43–46]. The electrophilic theory of carcinogenic chemicals introduced by James and Elizabeth Miller [11,47] was the first step in rationalization of the mode of action of animal carcinogens known by the 1970s. The Miller's hypothesis also helped to justify mutagenicity of chemicals towards Salmonella [48]. The electrophilic hypothesis has become a general theory of the carcinogens. The epigenetic carcinogens do not bind covalently to DNA and cause carcinogenicity through a large variety of mechanisms, while the genotoxic carcinogens are either electrophiles or can be activated to electrophilic reactive intermediates. During the last decade, several chemical functional groups or SAs have been identified for genotoxic carcinogens, based on Miller's theory. The identification of nongenotoxic carcinogens is much more challenging because there is no unifying theory for the explanation of their mechanisms of action.

John Ashby in 1985 introduced a list of SAs for carcinogenicity. This list contained eighteen SAs. The revised list of these SA can be found in Ashby and Tennant [48]. Each SA in the Ashby list has its specific mechanism of action. It is noticeable that there are some physico-chemical factors that may override the effect of these SAs in a molecule. The biological activity of a

molecule depends on different factors such as molecular weight, physical state, solubility and chemical reactivity. The Ashby and Tennant preliminary lists of carcinogenic SA was one of the most useful schemes to assess carcinogenic potential of substances with unknown carcinogenic properties. In 1996 Munro et al. created a functional groups list for genotoxicity based on the SAs of Ashby [49]. Cheeseman et al. [50] identified SAs useful to support higher threshold levels by using (Q)SAR, genotoxicity and short-term toxicity data. The identified SA were similar to the Ashby and Tennant list and were correlated with the TD50. The new list contained eight new more complex SAs. The list of SAs proposed by Ashby and Tennant and Cheeseman was revised by Kroes et al. [51].

Kazius et al. [46] expanded and refined Ashby's SAs by applying modern data mining techniques on chemical data of mutagenicity in Salmonella. Kazius et al. introduced a final set of 29 toxicophores which was able to classify the chemicals in the evaluation data set with 18% of classification error.

One of the most recent rule sets defined by human expert for mutagenic carcinogenicity has been developed by Benigni and Bossa [39,43]. The updated version of this rule set [39] is implemented in Toxtree version 2.6.13 [52].

1.9 Current Hazard Identification Procedures and Related Considerations

Since the carcinogenicity is a complex process, the rodent bioassay results are insufficient for accurate human health risk assessments [53]. Currently, genotoxic properties of the new chemicals are evaluated mostly by short-term studies. The chemicals that show genotoxic effects do not undergo long-term studies. Extrapolation of the results to humans from the rodent data is possible by considering the similarities and dissimilarities of the species.

Most of the identified carcinogens in humans are genotoxic or interact directly with DNA. Genotoxic chemicals are essentially different from nongenotoxic chemicals [16,54]. The classification of a carcinogen into genotoxic or nongenotoxic category has an important effect in the choice of the further studies and the indication of the chemical for risk assessment. Genotoxicity screening of a chemical by its metabolic information is more standardized compared to the chemical evaluation approach for defining nongenotoxic chemicals. While for evaluation of genotoxic chemicals a standard decision approach is approved [55], more effort is

needed for defining nongenotoxic chemicals [54]. Following the scheme of genotoxicity assessment reduces the number of animals used in carcinogenicity tests early in the evaluation process.

1.10 (Q)SAR and REACH

(Q)SARs are mathematical models that correlate the physico-chemical properties of chemicals to their biological activity such as toxicological and environmental fate properties. The (Q)SAR models are mainly statistical correlations, which describe a relationship between one or more quantitative characteristics of a chemical (descriptor) which is calculated from the chemical structure to a quantitative measure of property or activity of that chemical. These biological activities or properties for which the (Q)SAR models seek to estimate a predicted value are usually toxicological endpoints for human or environment. These prediction models can yield either continuous or categorical endpoint [56].

In other words, the information on the chemical structure of chemicals is connected to a specific property such as toxicity by means of mathematical and statistical methodologies and this relationship can be used as a predictive model for a new substance. Chemical applicability domain of each model should be defined by effective validation to make the model reliable for the new predictions. Considering the established applicability domain of each model, the reliability of the prediction of a substance by the model is decided. The chemically induced adverse effect of chemicals can be predicted by (Q)SAR models, as these models are becoming more and more robust and reliable. In addition, these models are fast and cost-effective and can replace a significant number of tests on animals and cells. The legislation on Registration, Evaluation, Authorisation and restriction of Chemicals, REACH, promotes the use of (Q)SAR models provided that, their scientific validity has been established, the substance falls within the applicability domain, the results are adequate for classification and labelling and/or risk assessment, and adequate and reliable documentation of the method is given. The REACH guidance has not determined any fixed criteria for the acceptance of the (Q)SAR models. In case a chemical is registered by an industrial registrant using a (Q)SAR model, the (Q)SAR model must be explained by them [56]. The application of (Q)SAR predictions can be useful in numerous fields, for example, in the initial phase of selecting chemicals for testing, experimental design of

experimental tests, evaluate and improve the data obtained from experiments, classification and labelling, and persistent, bioaccumulative and toxic assessment [56].

Another important field in which (Q)SAR predictions play role is in classification, clustering and read-across. Meaningful groups of chemicals can be created by the help of (Q)SARs. (Q)SARs are mathematical models that reveal the physico-chemical properties of the chemicals associated with their biological properties or activities. Presumably, these relationships modulate the activity giving rise to a trend development over a congeneric series of chemicals. The (Q)SAR predictions for molecular and toxicokinetic endpoints provide information for grouping and read-across process. Although, some (Q)SAR models may not provide adequate information for REACH or EU regulation CLP about the classification or the risk assessment of a compound or mixture, they can be used in a WOE approach together with other sources of data for designing a testing strategy and filling data gaps about the chemical properties [57]. According to REACH all other alternative testing options, such as (Q)SARs, should be considered before performing or requiring vertebrate testing [58]. All existing information on physico-chemical properties, toxicological and ecotoxicological data from *in vivo* and *in vitro* experiments and other non-testing methods must be gathered and put together for this end. Adding (Q)SAR and other non-testing methods, makes the information sufficient for the REACH requirements for the low tonnage substances. The REACH endpoint guidance claims that currently not all the mechanisms associated with reproductive toxicity can be identified by (Q)SAR models [59]. Although REACH demands and encourages the use of non-testing and *in vitro* methods to avoid vertebrate animal testing, unaccompanied (Q)SAR models do not produce reliable results that can replace whole-animal reproductive toxicity testing [59]. Supporting results from other experiments is needed to complete the negative result of the (Q)SAR and non-testing predictions for reproductive hazard assessment of a chemical. However, results of predictions of (Q)SAR models are useful in a WOE approach for grouping and read-across models and they contribute to reduction of animal tests [60].

According to the REACH Annexes VII-X for known genotoxic carcinogens or germ cell mutagens for which sufficient risk management measures are accomplished, no testing is

needed for reproductive effects [58]. Thus, according to the REACH guideline for chemical safety assessment [59], (Q)SAR results may contribute to reducing testing for reproductive toxicity. Among these points non-testing information which involves Quantitative Structure Property Relationships (QSPRs) and read-across can be used in accordance with the limitations explained for each individual endpoint. Each QSPR model has been built using a training set of substances and is more applicable to the chemicals which most closely match the samples used in the models. Therefore, the estimation of the QSPR models requires expert judgment. The predictions of such models need to be reasonable.

1.11 REACH Guidelines

Endpoint specific guidance of the REACH regulation [59] describes in what manner the WOE approach could be used for each endpoint. This section describes how the information collected from different sources could be integrated and used so that the conclusion on this information is sufficient for regulatory purposes (i.e. risk assessment). In other words, before proposing additional animal testing, use of alternative methods and adequacy of methods for generating additional information must be considered. It is precisely emphasized that experiments on vertebrate animals should be limited to the cases that all other data sources have been exhausted [61].

There are a number of issues determined by REACH to be considered before taking decision to perform the testing. These issues help to design fit for purpose *in vivo* tests, also provide evidence for not performing *in vivo* testing under certain circumstances.

- Testing requirements;
- Exposure/use pattern (emissions, yes or no, consumer use, etc.);
- Occurrence (monitoring data);
- Indications of the effect/ property based on animal or human data, *in vitro* data and non-testing information;
- Any concern e.g. based on toxicokinetics, read-across and (Q)SAR considerations,
- WOE;
- Seriousness of the effect;

- Other effects of relevance for the endpoint.[62]

1.12 Genotoxicity Assessments in the EU Regulations

The EURL ECVAM defines 3 endpoints that need to be assessed in the process of genotoxicity for the safety assessment of chemicals and the protection of human health and environment [63]. These 3 endpoints are: gene mutation, structural chromosome aberrations, and numerical chromosome aberrations (Table 2-page 25). Classification and labelling (C&L) of chemical substances is based on the results of the genotoxicity tests of the scientific tests for toxicity assessment in the EU [34] and in the world (UN GHS).

According to the EU legislations and directives there are two different approaches for the assessment of genotoxicity to humans. According to the first approach, a chemical that is nongenotoxic in the *in vitro* analysis is not considered for the further *in vivo* assessments (e.g. REACH, CLP and Cosmetics Directive). The second approach, foresees *in vitro* tests of chemicals followed up by *in vivo* assessments (e.g. ICH for pharmaceuticals and VICH for veterinary drugs). This decision is made because the alternative methods to *in vivo* test for carcinogenicity cannot thoroughly replace the animal tests [64]. Analyzing the regulatory requirements, EURL ECVAM suggests an efficient approach to improve the traditional genotoxicity assessment. The new approach has the objective of reducing the animal use in genotoxicity testing, and EURL ECVAM suggests that efforts should be directed towards the improvement of the current assessments while reducing the use of animals and at the same time satisfies regulatory information approach. The identified solutions for improving the genotoxicity assessment have the following aims:

- Increasing the performance of *in vitro* tests in order to avoid the additional follow-up *in vivo* tests;
- Improving the accuracy and quality of the *in vivo* follow-up testing to reduce unnecessary use of animals.

According to the chemical authorities the positive results of the *in vitro* genotoxic predictions need to be verified by *in vivo* tests, and this fact highlights the importance of finding solutions for reduction and refinement of genotoxicity tests. A strategy to reduce animal tests for

decision-making about carcinogenic or toxic compounds, can be collecting relevant data and drawing a conclusion on the basis of a data obtained from different sources. Pfuhler et al. [65] published some solutions to reduce animal tests in an ECVAM workshop report.

Table 2. Test methods most commonly used for genotoxicity/mutagenicity testing [63]

Test Method	**COUNCIL REGULATION (EC) No 440/2008 Test Method**	**OECD Test Guideline**	**endpoint**	**In vitro/ in vivo**
Bacterial reverse mutation test (Ames test)	B.13-14	TG 471	Gene mutations	vitro
In vitro mammalian chromosome aberration test	B.10	Updated TG 473	Structural aberrations	vitro
Mammalian cell gene mutation test	B.17	TG 476 (under revision)	Gene mutations	vitro
In vitro mammalian cell micronucleus test		Updated TG 487	Structural and numerical aberrations	vitro
Mammalian erythrocyte micronucleus test	B.12	Updated TG 474	Structural and numerical aberrations	vivo
Mammalian bone marrow chromosome aberration test	B.11	Updated TG 475	Structural aberrations	vivo
Transgenic rodent somatic and germ cell gene mutation assays		TG 488	Gene mutations	vivo
In vivo mammalian alkaline comet assay		TG	DNA damage	vivo

1.13 Read-across

A read-across approach finds out the relevance or relationship between the properties of the chemical structures and then make assessment on the applicability of this information to another substance. It is crucial to detail the reasoning behind the inference on the substance for which the property is unknown.

A read-across process which is based on the concept of similarity can be applied in different forms: one-to-one (a similar substance can be used to make an estimation for a target substance) b) many-to-one (two or more analogues used to make a prediction for a single substance c) one-to many (one analogue used to make estimations for two or more substances) d) many-to-many (two or more similar compounds used to make estimations for two or more substances).

There are some important issues to be considered when using a read-across model. These characteristics are as follows:

- The source substances must have the same structural features and the functional groups of the target substance;
- the physico-chemical profile of the similar compounds must be comparable to those present in the target substance;
- the relevant molecular descriptors must have comparable values;
- the analogue substances must have approximately the same molecular weight.[66]

The results of a read-across should be interpreted using expert judgement and for the support of the conclusion detailed documentation is required. The read-across approach is more suitable for the physical hazard related to physico-chemical properties of the substances, as reliable test data should be available according to the CLP regulation. Therefore, if read-across is used as a unique method to generate a value to meet the endpoint data requirements, the criteria given in section 1.5 of Annex XI to REACH must be met.[66]

1.14 Classification and Labelling and Chemical Safety Assessment

Knowledge about physico-chemical properties of the chemicals and chemical safety assessment is important for the environment and human health. All the stages of the substances' lifecycle must be assessed and controlled in the process of chemical safety assessment, these stages include manufacture, transfer, use and disposal of the chemical substances. Further, physico-chemical data are essential for the correct planning of (eco)toxicological studies and for the optimization of the test conditions.

The standard test and most confident assay for carcinogenicity is the 2-year rodent carcinogenicity bioassay determined and described by OECD. The purpose of this assay is "to observe test animals for a major portion of their life span for the development of neoplastic lesions during or after exposure to various doses of a test substance by an appropriate route of administration." Usually two species (mice and rats) and both sexes are used in this test. The chemical exposure is dosed and executed by oral, dermal or inhalation modes based on the expected human exposure. Dosing is done during two years and animal health is screened

throughout the test. The most important results of the test is obtained by the thorough examination of the animal tissues and organs at the termination of the assay.

The combination of carcinogenicity and chronic toxicity animal bioassays endpoints may reduce the animal use [67]. A number of transgenic rodent models have been suggested as alternatives to the standard bioassay carcinogenicity test by the ILSI HESI, but none of them was as efficient as the traditional 2-year assay for identification of carcinogens [68]. Most of these models were capable of detecting genotoxicity that can be already detected by other *in vitro* genotoxicity assays. Alternative models are not still suitable for the detection of nongenotoxic carcinogens [69].

1.15 Development and Optimisation of Alternative Methods

1.15.1 Importance of Mode of Action and Weight of Evidence Approach

The most appropriate key events are needed to understand the mechanisms of action of nongenotoxic carcinogens. Since there exist numerous modes of action for these substances, a WOE approach seems to be essential to deduce a reasonable conclusion out of the gathered data for a chemical.

There are nongenotoxic carcinogens in IARC group 3 (i.e. not classifiable as to its carcinogenicity to humans) which are not carcinogenic in humans. To evaluate these group of carcinogens, the WOE approach is a useful tool. This approach helps the scientists to understand the differences of modes of action in rodents and humans.

1.15.2 Alternative Methods for Detecting Nongenotoxic Carcinogens

In the process of the detection of the nongenotoxic carcinogens and exploring alternative methods for their assessment, it is important to consider the vast range of modes of action of these chemicals. These modes of action include: mitogenic induction, inhibition of gap-junctional intercellular communications, endocrine modifiers, oxidative stress, immunosuppressants, regenerative proliferation and/or DNA methylation. Some of the examples of alternative methods for the nongenotoxicity detection are: (Q)SARs, measuring replicative DNA synthesis as an indication of cell proliferation, the *in vitro* cell transformation assays, measurements of inhibition of gap junction intercellular communication [70] and the use of

gene expression profiles with mechanistic networks for the identification of potential markers of nongenotoxic carcinogens.

1.16 Quantitative Structure–Activity Relationship (QSAR)

The main results of *in vitro* cell toxicity used by (Q)SAR models for nongenotoxic carcinogenicity are: several markers of *in vitro* cell toxicity including inhibition of gap-junctional intercellular communications, modulation of apoptosis and induction of cellular proliferation [71]. In this study, the structural features of the nongenotoxic carcinogen associated with toxicity or ligand binding, as in the case of estrogen, peroxisome proliferators and tubulin protein receptors, have been analyzed [71].

The (Q)SAR models for detection of carcinogenicity use information and correlate biological activity or chemical reactivity to chemical structure. These models are practically based on the assumption that similar chemicals have similar activities. The OECD has defined a number of principles for the validation of the (Q)SAR models for regulatory analysis [72,73]. These principles for (Q)SAR models are: having a defined endpoint, an unambiguous algorithm, and measures of goodness-of-fit, robustness, predictivity, and applicability domain.

A great number of nongenotoxic carcinogens are mutagenic inducers which cause cancer by increasing cellular proliferation. Hepatocyte rodent *in vivo* studies indicate that most of the hepatocarcinogens cause cancer by accelerating hepatocyte division [74–77].

As a result of putting into practice the REACH legislation, the number of 2-year rodent carcinogenicity bioassay is reduced and this fact can lead to the lack of detection of a large number of nongenotoxicity carcinogens. Because of the high risk of hazard associated with this group of carcinogens, there is an increasing need of alternative methods for their detection. Possible alternative methods for this purpose include: SARs and (Q)SARs, replicative DNA synthesis assay, the *in vitro* cell transformation assay and/or inhibition of GJICs. None of these alternative methods provide any information on the mode of action, thus further studies are needed to fill this gap of data. Using toxicogenomics to analyse multiple pathway-specific gene expression profiling is an efficient method to identify putative alerts. Additionally, statistical validation studies to examine the sensitivity, specificity and accuracy of these models play an

important role in improving these alternative models. It is important to discover as much as possible different modes of action of nongenotoxic carcinogens and not to depend only on the traditional nongenotoxic carcinogens identification methods (tetrachlorodibenzo-p-dioxin, carbon tetrachloride, and cyclosporine).

1.17 Software Packages for Mutagenicity and Carcinogenicity Predictions

Two major alternatives to *in vivo* testing are *in vitro* and *in silico* techniques. In the last decade, numerous computer software have been developed in order to replace, reduce and refine the animal tests. These software packages include also mutagenicity and carcinogenicity predicting models of the chemical compounds.

1.17.1 VEGA Platform

The VEGA platform contains a number of (Q)SAR models for predicting mutagenicity and carcinogenicity such as CAESAR. Two new carcinogenicity models have been added to the VEGA platform by the author [78]. CAESAR ((Q)SAR mutagenicity models) was specifically developed for the REACH regulation in collaboration with the United States Environmental Protection Agency (http://www.caesar-project.eu/). The mutagenicity models in the VEGA platform are based on data obtained from the Ames bacterial test. Models on carcinogenicity, developmental toxicity and etc. are freely available from the VEGA platform.

1.17.2 DEREK Nexus

The DEREK Nexus [79] is a knowledge-based expert system, developed by LHASA Limited that predicts the genotoxicity, mutagenicity and carcinogenicity of a chemical by highlighting the SAs present in its molecular structure. Derek Nexus toxicity predictions are a result of two processes: evaluating SAs and estimating the likelihood of toxicity.

The knowledge-based SAs for *in vitro* mutagenicity have been implemented by experts who have assessed relevant Ames data and supporting mechanistic data (e.g. DNA adduct formation experiments). If a query compound matches a SA, the alert will fire with an associated reasoning level (e.g. plausible, probable or certain). The reasoning levels associated with the *in vitro* bacterial mutagenicity alerts in Derek Nexus gives an indication of the likelihood for compounds in that class to be active in the Ames test [80].

Each bacterial, *in vitro* mutagenicity alert in the knowledge base was examined by a scientist with expertise in mutagenicity alert development. The patterns encoding the SAR for each alert were modified using the Derek Knowledge Editor if they contained features that were implemented to prevent the pattern being activated by nonmutagenic compounds (so-called exclusion patterns). Such features were removed and the resultant 'predictive space' was stored within a modified knowledge base. Thus, each bacterial, *in vitro* mutagenicity alert in Derek had a corresponding region of predictive space [81].

1.17.3 TOPKAT

TOPKAT [82] is a (Q)SAR-based system, developed by Accelrys Inc. (http://accelrys.co m/). Some of the TOPKAT toxicological endpoints are mutagenicity, developmental toxicity, rodent carcinogenicity, rat chronic LOAEL, rat MTD and rat oral LD50. TOPKAT models are developed using two-dimensional molecular, electronic and spatial descriptors. The toxicity prediction is obtained from a chemical's molecular structure. TOPKAT defines an applicability domain value which estimates the confidence in the prediction by applying the patented Optimal Predictive Space validation method. Any prediction generated for a query structure outside of the OPS space is considered unreliable.

1.17.4 MultiCASE

MultiCASE [83] (MultiCASE Inc., Cleveland, OH, USA) is a prediction model for genotoxicity and carcinogenicity endpoints based on US FDA and US EPA. MultiCASE identifies SAs with a potential to initiate high biological activity, in addition, some statistical parameters are analysed to complete the predictions. The mutagenicity and genotoxicity models are based on the data obtained from Ames test, direct mutagenicity, base-pair mutagenicity, frameshift mutagenicity, chromosomal aberrations, and sister chromatid exchange data. The carcinogenicity model includes different rodent assays (rate, mouse, male, female, and TD50 rats) and human epigenetic studies. All models use the statistical approach with the exception of the rule-based model for the Ames mutagenicity.

1.17.5 QSAR Toolbox

QSAR Toolbox [84] in cooperation with the ECHA is a read-across tool for grouping the chemicals and examining their toxicity effects according to the OECD principles

(http://www.qsartoolbox.org/). QSAR Toolbox systematically groups chemicals into classes according to their molecular structure, physico-chemical and biological properties. This software extracts structural characteristics and modes of action based on experimental information for the target molecule. The common mechanisms of action and common toxicological behaviour or consistent trends among results related to regulatory endpoints are results for an evaluation in this prediction software.

1.17.6 Toxtree

Toxtree [52] is a free tool for the assessment of mutagenicity and carcinogenicity of the chemicals using decision trees. Toxtree mutagenicity and carcinogenicity model is based on the SAs of the Benigni-Bossa rule set, SAs for identification of Michael acceptors, and SAs confirmed by positive *in vivo* micronucleus tests. The program identifies any SA present in the target molecule structure and concludes about the mutagenic or carcinogenic property of the chemical compound under investigation. The result of the prediction can be class I (inactive), class II (weak activity), or class III (active).

1.17.7 LAZAR

LAZAR [85] is an open source tool for the prediction of carcinogenicity and Salmonella mutagenicity. LAZAR creates local endpoint (Q)SAR models based on a training set (only nearest neighbours) for each chemical separately. It first calculated the descriptors and determines the molecular similarity and then it builds a local (Q)SAR model based on a database of experimental toxicity data. This program meets all five OECD principles.

1.17.8 ACD/Tox Suite

The ACD/Tox Suite [86] package contains predicting models for genotoxicity and carcinogenicity. The assessments are made based on validated (Q)SAR models in combination with expert knowledge. The software highlights and identifies the SAs which are responsible for toxic properties and extracts some similar molecules from the training set. The training set is composed of compounds that are genotoxic in Ames test.

1.17.9 Leadscope Model Applier

The Model Applier developed by Leadscope [87] (Leadscope Inc., Colombia, OH, USA), uses (Q)SAR models for Salmonella mutagenicity, E. coli mutagenicity, mouse lymphoma, *in vitro* chromosome aberrations, and *in vivo* micronuclei.

1.17.10 SARpy

SARpy [88] is a data mining tool which works in a SAR approach. This tool is able to identify a list of active and inactive molecular fragments that act as SAs for biological activity from a learning set of chemical compounds with known binary classification toxicity property (e.g. carcinogenicity, mutagenicity, etc.). The entire process is designed to fit with human reasoning. In fact, SARpy is a computerized tool which helps the expert to extract significant molecular substructures with potential effect in toxicity or in nontoxicity in an automatic way with customized requirements to be set by the user. The extracted list of active and/or inactive SAs identified by SARpy are considered new prediction models with satisfactory prediction ability on an external test set, in particular in the case of mutagenicity [89].

1.18 Weight of Evidence

Human health and environmental decision-making is often based on multiple lines of evidence. WOE is a process for integrating different and sometimes conflicting sources of information (lines of evidence) to determine a relative support for possible answers to a scientific question or assessment. A line of evidence is a group of evidence of similar type which pertain to an important aspect of the environmental or human health assessment. The distinct elements of information forming a line of evidence are called “studies” or “pieces of evidence”. The multiple sets of information of lines of evidence can be divided into three types: the biological field line, the toxicity line and the chemistry line [90]. Combining information from multiple sources to be used in decision-making is not a simple procedure. From a regulatory point of view, risk assessors make use of WOE approaches to perform integration and reach conclusions [91] in a qualitative or quantitative manner. Also the industry employs different sorts of WOE approaches in the toxicity assessment of the chemicals [92]. The guidance on WOE provided by EFSA introduces a general framework of approaches used to weigh the lines of evidence in

order to find an answer for any scientific question which needs to consider different sources of information and integrate various data in its assessment process. This document explains types of qualitative and quantitative approaches in the WOE field and lists the relative methodologies. According to the EFSA document, the WOE assessment consists of three steps:

1. Assembling the evidence
2. Weighing the evidence
3. Integrating the evidence

The first step compromises searching and selecting relevant evidence for answering the question, and also grouping the evidences found into lines of evidence. The second step involves the evaluation and assigning weight to the evidence. In the third step, the collected lines of evidence are integrated to reach conclusions, by weighing the relative support for possible answers to the question at hand.

During the process of a scientific assessment the three above mentioned steps may be required at one or more points, whenever integration of evidence is necessary to reach a conclusion. Problem formulation is the first stage of scientific assessment in which the question to be addressed by each WOE assessment is defined. The three steps of the WOE framework is the preceding step in the scientific assessment. The outcome of the process of WOE influences directly or indirectly the overall conclusion of the scientific assessment. The existing uncertainties that may affect the overall assessment are evaluated during the steps of the WOE and in addition, a separate step is considered for the analysis of these uncertainties as the last stage of the scientific assessment before any conclusion is reached. In some assessments an additional sensitivity and influence analysis is performed to examine the influence of evidence and uncertainties that may influence the conclusion. This process is iterating, and permits the assessor to return to an earlier step in order to refine the scientific assessment.

Reliability, relevance and consistency are mentioned as three key considerations for weighing evidence in many scientific publications. The quality of the evidence considered for supporting an answer to a scientific question (reliability), how applicable the evidence is to that question (relevance) and how consistent the line of evidence (or the piece of evidence) is with other existing evidence to answer the same question (consistency). Relevance needs to be considered

when identifying evidence, and relevance and reliability are the key considerations when selecting which evidence to be concluded in the assessment. Relevance and reliability may be different in the selected evidence and this will be analyzed in the phase of weighing the evidence.

Reliability measures the correctness of a piece of evidence, and whether this piece of evidence represents correctly the quantity, property or event it refers to. Reliability consists of accuracy and precision.

Relevance defines to what extent a piece of evidence is relevant to answer a specified question, provided that the information it consists is reliable.

Consistency explains the compatibility of the information obtained from different pieces of evidence, after analysis of reliability and relevance.

The important role of multi-criteria decision analysis and WOE methodologies in environmental decision-making [93,94], as well as, hazard ranking of NMs [85,95–97] has been assessed by numerous scientists.

1.18.1 Weight of Evidence Method Classification

Linkov et al. [98] provided a brief review of qualitative and quantitative WOE approaches in human and environmental risk assessment. Linkov et al. introduced a classification system for characterizing WOE methods, based on Weed [99] and Chapman et al.'s [100] (Table 3-page 35) The methods implied in this table are ordered from the most qualitative methods (e.g. listing evidence and BPJ or narrative review) to the most quantitative methods (e.g. indexing and quantification) which incorporate methods involving decision analysis tools and defining the problem as statistical hypothesis testing.

Table 3. Weight of evidence methods [98]

Method	Method description
Listing Evidence	Presentation of individual lines of evidence without attempt at integration
Best Professional Judgment	Qualitative integration of multiple lines of evidence
Casual Criteria	A criteria-based methodology for determining cause and effect relationships
Logic	Standardized evaluation of individual lines of evidence based on qualitative logic models
Scoring	Quantitative integration of multiple lines of evidence using simple weighting or ranking
Indexing	Integration of lines of evidence into a single measure based on empirical models
Quantification	Integrated assessment using formal decision analysis and statistical methods

Listing evidence as the simplest form of WOE, collects lines of evidence together. Lines of evidence are presented without any integration phase. The assessors may reach a conclusion considering the list of evidence. All the other methods presented in Table 3 (page 35) include a form of integration. BPJ is similar to listing evidence, but it attempts to integrate the evidence by invoking a professional opinion. Casual criteria methods consist of methods evaluating cause and effect relationships. Casual criteria and logic provide a consistent structure for the analysis and transferability of the methods are improved by these methods. Casual criteria methods by means of outlined criteria, establish a cause and effect criteria, by illustrating that the specific criteria is met. While logic methods make use of previously outlined methods for integrating lines of evidence, such as US EPA carcinogenicity guidelines [101]. Casual criteria and logic leads to a more transparent integration, but the methods are qualitative and may be biased by experience. Casual criteria and logic depend on BPJ to synthesize lines of evidence. Scoring is the simplest quantitative method of WOE which assigns weights to lines of evidence. A numerical WOE score is a combination of weights calculated by different methods such as BPJ

based on consistency, specificity and strength of the association. Indexing determines a single value as the outcome of the analysis by integrating all the weights assigned to each line of evidence. Scoring and indexing do not quantify judgments using formal decision analysis or probabilistic techniques. Formalized mathematical methods involving quantitative methods are used by quantification category to weigh the evidence. Non linearity and correlations can be integrated in the methodologies of quantification category. Also in this category scientific results with individual expert or decision maker judgment and comparison across multiple experts can be integrated in a transparent and reproducible manner. Multiple-criteria decision analysis is a quantification method which uses likelihoods to weigh the evidence.

1.18.2 Weight of Evidence Application

Linkov et al.[98] reported a summary of methods and application areas of WOE. The WOE applications are divided into two main categories: human health and ecological. Under category of human health that we focus on, WOE can be applied to i) method development, which develops methodologies for human health risk assessment, ii) toxicity analysis, which assesses the adverse effects of a substance, iii) mode of action determination, the modes in which a substance may cause harm to human health, iv) benchmark development, defines the allowed exposure levels of various substances, and other fields. This literature review yielded a comparable number of human health and ecological uses. Human health methodologies are performed mostly by best judgment methods, and a minor number of studies are conducted by quantitative methodologies. The qualitative methods are based on BPJ as recommended by the US EPA guidelines. Casual criteria and logic are mainly used by ecological methodologies. Overall, this review indicates that qualitative methods of WOE are used more than quantitative ones in the applications of WOE and BPJ is the most widely used method.

1.18.3 Weight of Evidence Approach in Nanomaterials Risk Assessment

Hazard identification is an important step in the process of NMs risk assessment and is required under regulatory frameworks of the US, Europe and worldwide. The current risk assessment methods used for chemical and biological materials may not be adoptable for NMs because of the existing uncertainties in identifying the relevant physico-chemical and biological properties

that are able to adequately describe the NMs. Understanding and managing the impact of NMs on human health and environment need new approaches. Currently, our knowledge about the toxicity of NMs is barely comprehensive [102,103]. The potential toxicity as well as potential for exposure and risk of NMs may be highly impacted by their physico-chemical characteristics. In the absence of definitive data, NM research and regulations can make use of a systematic characterization of factors to identify toxicity and risk of NMs [104]. On the other hand, given the complexity of NMs and high uncertainties associated with them, multiple studies on the characteristics and potential hazard effects result in varying data points. WOE approaches are recommended for NMs risk assessment for prioritizing research studies and identifying NMs with hazard effects. Expert opinion is frequently used to fill the knowledge gaps during decision-making. Influence diagrams by using expert judgment approach are designed to assess the risk of NMs on human health [105]. Linkov et al. [96,97] proposed a WOE approach for NMs regulation and management. For the evaluation of the adverse effects of NMs, different aspects such as life cycle and characteristics of NMs must be addressed. To follow the European Commission PEC/PNEC (the ratio of predicted environmental concentration (PEC) and predicted no effect level (PNEC)) approach of the TGD [106], there is lack of information about toxicity and characteristics of NMs in order to estimate exposure concentration and no effect concentration and thus the risks. Available information about the life cycle phases of NMs concerning different issues such as environmental behavior relevant for exposure and effect are gathered in lines of evidence. Finally, the information gathered in these lines of evidence are integrated to reach a conclusion about the adverse effects of NMs. The results are usually qualitative, assigning the assessed NM to a specific class. This ranking system helps the assessors to prioritize the most potential NMs which may pose risk to human health or environment.

Zuin et al. [95] reports the application of a WOE procedure to the assessment of NMs that may cause harm to human health. The procedure is divided into three steps.

The starting point in the assessment of NMs is physico-chemical characterization and properties and toxicological information. Physico-chemical properties of NMs are the first line of evidence, which is mainly related to the capacity to evaluate the potential exposure according to

their physico-chemical properties, such as adsorption tendency and bioaccumulation potential. The second line of evidence addresses the evidence that a NM may enter the body and cause adverse biological harms. For the estimation of the hazard effects, within each line of evidence other indicators, such measurement endpoints are identified. These indicators should include a wide range of information, like physico-chemical properties, toxicological endpoints, and data obtained from literature.

The second step is defining rating classes for each indicator. The assignment of each NM to a specific rating class may be performed on the basis of (i) characterization activity and toxicity test, (ii) expert judgment, (iii) literature data. The defined classes can be high, moderate, low and negligible. The qualitative and quantitative values associated with the indicators permits the NM to be assigned to a class, on the basis of its properties-related exposure level and toxicity.

Ranking procedure is the last step of the process of the WOE approach-based method. In the suggested procedure the calculation of frequency of hazard occurrence is performed on the basis of the rating class assigned to different indicators.

1.19 Abstracts of the Main Parts

1.19.1 Part 1- New Clues on Carcinogenicity-Related Substructures Derived From Mining Two Large Datasets of Chemical Compounds

The first section of my thesis is about development of two SA-based (Q)SAR models for carcinogenicity. Two (Q)SAR models were developed by extracting data from well known carcinogenicity databases with genotoxic and nongenotoxic carcinogenicity reliable data gathered from rodent *in vivo* tests, Ames test and *in vitro* tests and epidemiological data. (ANATARES [107] carcinogenicity database and the combination of Kirkland et al. [108] and ISSCAN [109] database).

1.19.2 Part 2 - Toxdelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances.

The second part explains a new read-across tool embedded in the ToxRead software: ToxDelta. Two structurally similar molecules share an MCS. In order to evaluate if two similar molecules have different effects, we focused our attention on the molecular fragments which are not in the MCS. These parts may increase or decrease the value of the property. We considered a variation of the MCS concept of efficient relevance in toxicity assessment where the rings of molecules must not be broken. To assess the toxicity of the target chemical, ToxDelta extracts the MCS and delineates the remaining fragments. Each of these fragments moiety represents a difference between two molecules and its relevance in the toxicity assessment are evaluated against a knowledge based list of active and inactive fragments. ToxDelta considers the dissimilarities of the molecules in a read-across approach.

1.19.3 Part 3 – Genotoxicity Induced by Metal Oxide Nanoparticles: a Weight of Evidence Study and Effect of Particle Surface and Electronic Properties

Genotoxicity of metal oxide NMs is an endpoint with intensive testing resources mainly resulting from *in vitro* comet assay. Current contributions to the genotoxicity data assessed by the comet assay provide a case-by-case evaluation of different types of metal oxides that ranged from 15 to 90 nanometres and had different crystal structures. We have assessed the quality of a multi-source data set of *in vitro* comet assay data retrieved from genotoxicity profiles for 16

bare nano-sized metal oxides with different chemical core compositions. There is an inconsistency in the literature as to the genotoxicity testing data that requires intelligent strategies, such as WOE evaluation. An evaluation criterion was applied to establish which of these meta data were of sufficient quality and what weight could have been given to them in inferring genotoxic results. We surveyed the collected data on 1) minimum necessary characterization of NPs, and 2) principals of correct comet assay testing for NMs. We quantum-chemically calculated a set of structural descriptors for the 16 metal oxides. A classification model based on a decision tree has been developed for the prepared dataset. Three descriptors have been identified as the most relevant variables to genotoxicity prediction in our classification model: heat of formation, molecular weight and area of the oxide cluster based on conductor like screening model. The proposed genotoxicity assessment strategy that is based on quantum-chemical descriptors is useful to prioritise the study of the NMs, which may lead to high risk for human health.

CHAPTER 2

2. Aim of the Study

Human health and environmental risk assessment draw upon multiple sources of information. Combination and integration of multiple lines of evidence is required to draw conclusions about the risks posed by chemical compounds and NPs. To this end, scientists and assessors take advantage of WOE approaches. The increasing need for new non-testing methods for the estimation of toxicity of chemicals coincides with the emergent necessity of more adequate WOE approaches to facilitate the process of decision-making. The proposed non-testing methods such as (Q)SAR and read-across are promising methodologies that help the assessors to fill the gaps of information, which in their turn provide means that can be used in risk assessment in a WOE approach. The regulatory bodies (such as EFSA) and industry already make use of WOE approaches in the process of decision-making.

The main aim of the present study is to introduce a new WOE framework for the results of *in silico* models, which explores the structural properties of the target and the similar compounds (used in read-across terms) and combines the estimations of the (Q)SAR models with the toxicity information associated with the molecular structures in an interpretable manner.

In this dissertation, for the improvement of the *in silico* ((Q)SAR and read-across) methods in genotoxicity assessment of chemicals (i) two new structure-based carcinogenicity models, (ii) a new read-across model based on maximum common substructure for mutagenicity endpoint, are developed. Additionally, (iii) a genotoxicity model for the metal oxide NPs is introduced.

We addressed two WOE methods in the present study. BPJ method as a qualitative approach is used in the study of genotoxicity of metal oxide NPs. Given the very high uncertainties with NMs and their hazard effect on environment and human health, research resources should be conducted in the direction of integrating the available data to help the manufactures, regulators, consumers and other stakeholders in the process of decision-making. A genotoxicity data base of metal oxide NPs is developed using a BPJ framework for selecting the most reliable data from peer reviews. A decision tree model is built for classification of these data base using three quantum-chemical descriptors, by the machine learning means.

Multiple studies for hazard identification of chemical substances often result in varying data pointing in different directions, and this causes conflict in the process of decision-making when data interpretation is attempted. The majority of WOE frameworks for the analysis of the results of *in silico* models are qualitative and do not satisfy the growing need of objectivity and transparency that are essential for regulatory purposes. Most often, the prediction results of (Q)SAR or read-across models lack the reasoning part, by which the assessor can interpret and clarify the objectivity of the hazard predictions for a chemical under investigation. The structure-based models such as the two models introduced in the first part of the present document, are adequate means to resolve the shortcomings of the *in silico* models with no interpretation for their estimations. These structure-based models are also useful for the analysis of the similar compounds to address the structural similarities and dissimilarities and the presence or absence of the known SAs in individual chemical during the process of comparison and inference. ToxDelta as a new read-across tool is developed for the comparison between the target and the source chemicals, and to analyse the different fragments and their role in amplifying or reducing the toxicity.

In addition, we discussed how to deal with different and sometimes conflicting data obtained from various *in silico* models ((Q)SAR and read-across) in qualitative WOE terms based on identification of SAs in a read-across approach to facilitate structured and transparent development of answers to scientific questions. The results of different (Q)SAR models are explored together with the analysis of the similarities and differences of the similar compounds in read-across terms. The study of the structural similarities and dissimilarities between the target and the source compound helps the expert to validate or revoke the assumption that the properties of the similar compounds can be assigned to the target compound. To show the utility of the use of multiple tools ((Q)SAR, ToxRead and ToxDelta) within an integrated WOE prospective to obtain a toxicity conclusion, the application of the framework is illustrated using two drugs as case studies: Valproic acid and Diclofenac.

2.1 Part 1 – Carcinogenicity Models

The identification of the SAs is a compelling factor in understanding mechanisms, and assessing the hazard risk of chemicals. The already known lists of carcinogenicity SAs, including genotoxic and nongenotoxic moieties can still be refined and enhanced by further studies on carcinogenic substances.

Within the ANTARES project (LIFE08 ENV/IT/000435), a carcinogenicity data set is developed for testing the performance of seven software packages. This dataset comprises 1543 chemicals together with their carcinogenicity values apparently of good quality. The carcinogenicity properties of the chemicals are related to rat toxicity (presence of carcinogenic effects in male or female rats). This dataset is a combination of the CAESAR data set and "FDA 2009 SAR Carcinogenicity - SAR Structures" database. The CAESAR data set encloses 805 chemicals taken from DSSTox Public Database Network (http://www.epa.gov/ncct/dsstox/sdf_cpdbas.html) which was developed from the Lois Gold's CPDB. All the chemicals have been cleaned and cross checked in this data set. ID number, chemical name, CASRN, experimental carcinogenic potency (TD50) values for rat and corresponding binary carcinogenicity classes are supplied for each chemical in this data set. In addition, 739 compounds not present in the CAESAR dataset were taken from the "FDA 2009 SAR Carcinogenicity - SAR Structures" database using the Leadscope software (http://www.leadscope.com/). A categorical label for carcinogenicity was already provided in the original database. A compound was labeled as carcinogenic if a positive outcome was detected in male or female rats.

The ANTARES rodent carcinogenicity data set is a promising source for developing models for classification of carcinogenic potency. Accurate preprocessing of data and selection of data with rats carcinogenicity provide consistent data suitable for QSAR modeling with carcinogenic response closer to human. Additionally, a wide diversity of molecular structures, thus a diverse number of chemical classes and biological mechanisms improves the carcinogenicity prediction ability of QSAR models.

Taking advantage of SARpy [88], as a statistical fragment extraction tool -without any a priori information- on a large data set of chemicals, is a promising strategy to inspect carcinogenicity

potential SAs. The SAs extracted by SARpy are subsequently revised and checked by human expert and a list of more significant SAs has been created out of the initial rulesets. The first carcinogenicity model is the collection of these SAs.

The second model is based on a combination of the ISSCAN database [109] and the Carcinogenicity Genotoxicity eXperience (CGX) database created by Kirkland et al. [108].

The ISSCAN database contains information on chemical compounds tested with the long-term carcinogenicity bioassay on rodents (rat, mouse). For the chemicals in the ISSCAN database this information is provided: carcinogenic potency in rat and mouse, mutagenicity in Salmonella typhimurium (Ames test), carcinogenicity results in the four experimental groups most commonly used for the cancer bioassay, carcinogenicity results from the NTP experimentation (when available), overall carcinogenicity, and the source of carcinogenicity data.

Kirkland et al. [108] used a battery of three commonly used *in vitro* genotoxicity tests—Ames + mouse lymphoma assay + *in vitro* micronucleus or chromosomal aberrations test— to classify rodent carcinogens and non-carcinogens, inside a large database of over 700 chemicals compiled from the CPDB, NTP, IARC and other publications. A WOE approach has been applied to integrate the results obtained for these chemicals from the literature. 940 chemicals present in the Kirkland et al.'s dataset were merged to the ISSCAN data set. The duplicates and the conflicting values have been eliminated. The resulting dataset is conventionally named ISSCAN-CGX data set which contains 986 chemicals together with their carcinogenicity calls. The ISSCAN-CGX data set contains human-based assessment data on carcinogenicity. The second carcinogenicity model is developed by extracting the active SAs from the ISSCAN-CGX data set by means of SARpy.

The advantage of developing QSAR models using such large datasets is that changes in "calls" for a small number of chemicals containing a potential SA will not significantly influence the overall findings. In addition, there is more chance to extract the potential carcinogenicity fragments from a data base which contains chemicals with more functional groups. A model that considers more molecular functional groups has a wider applicability domain. All the extracted rules or SAs are examined by an expert and their significance in carcinogenicity is verified based on his expertise.

It is to be noted that rodents are considered more sensitive than humans to carcinogen chemicals. The ANTARES model being developed on pure rodent data is supposed to be more conservative rather than the ISSCAN-CGX model which is developed using integrated data obtained from more carcinogenicity assays in a WOE approach by human expert.
The aim of the first part of my study was expanding and upgrading the knowledge on the SAs by the help of artificial intelligence and data mining approaches. The new carcinogenicity models have been implemented in the VEGA platform with notable prediction results. In both rule sets some new SA have been identified.

2.2 Part 2 – ToxDelta

The read-across approach is based on the similarity property principle. The similarity property principle states that structurally similar molecules are more likely to have resembling properties. Contrarily, studies show that structural similarity does not always imply similarity in activity [110] nor in descriptors [111]. Minor modifications can make active molecules to lose their activities completely. Structurally similar compounds can have very different properties. We contemplated the effects that the dissimilarities may trigger on the properties of the structurally similar compounds. To this end, we developed a new read-across tool, ToxDelta, to accomplish the effects of the dissimilarities in a read-across approach. This new tool is considered as a complementary tool to be implemented in the ToxRead software.
ToxRead, is a new read-across tool developed by our group. ToxRead is an ad hoc visualization and data search method which use similarity measures and SA search to organize in a chart a picture of all the relevant information. ToxRead, with its original representation of the read-across results makes it easy for the user to move in different directions of toxicity or nontoxicity properties of the target and source molecules which share the same SAs. ToxRead does not provide exclusively toxic SAs, it extracts and depicts also the nontoxic SAs present in the target molecule. ToxRead is currently applied to mutagenicity and bioconcentration.
At the present moment, the ToxDelta tool is a stand-alone software for the read-across mutagenicity assessment. ToxDelta focuses on the dissimilar substructures between two similar molecules, and analyses whether these dissimilarities reduce or amplify the toxicity in the target

chemical compound. The MCS between two structurally similar molecules has been already analysed by ToxRead, and all the possible SAs present in the sharing structure has been already examined. As a further assessment, ToxDelta takes a closer look at these fragments and exploits the role of these differences in the hazard risk of the chemicals. The characteristics of the extracted molecular fragments is assisted in an a priori list of mutagenicity SAs.

2.3 Part 3 – Metal Oxide NMs Genotoxicity Model

NMs are utilised in many fields of industry, medicine and military applications. The NMs' potential hazardous effects can cause a wide range of damage to human health and environment. While the acute toxicity of NMs has been addressed in many studies, genetic toxicity, in particular genotoxicity of the NMs still needs to be explored by more scientific works.

In order to examine the toxic effect of a single NM, given the diversity within each group of NM a large number of property combination need to be considered (different shapes, size, crystallography, etc.). The risk assessment in a case-by-case manner makes the task challenging. In addition, increasing the number of in vivo tests is opposing to the Russel and Burch's 3R principle to replace, reduce and refine animal testing of the EU Directive 2010/63/EU. The REACH regulation promote exploiting all existing data and focusing on new approaches e.g. non-testing methods and data integration using WOE strategies as effective tools to achieve this goal.

In this study, we assessed the genotoxicity properties of sixteen metal oxide NPs in a WOE approach using the results obtained by *in vitro* Comet assays. Different peer review studies from 1994 to 2014 about genotoxicity of metal oxide NMs using *in vitro* Comet assay have been collected. The reliability and the relevance of these scientific articles have been assessed taking advantage of the WOE technique. The overall assessment of a series of queries for this assessment led us to assign a genotoxic property to each of the metal oxide NM under study as a conclusion. In addition to the preparation of a list of metal oxide NMs with their genotoxic overall effect, we decided to study the relationships between quantum-chemical / physico-chemical descriptors of these NMs and the overall assessment of their genotoxicity effects. We quantum-chemically calculated a series of quantum-mechanical descriptors for the set of metal

oxide NMs in our dataset and then opted for an understandable classification method, considering the binary classification nature of the endpoint. In order to exploit any significance relationship between the quantum-chemical descriptors and genotoxicity of the prepared dataset, a tree decision model is applied to the dataset. Three descriptors have been identified as the most relating variables to genotoxicity property in our classification model: heat of formation, molecular weight and area of the oxide cluster based on conductor like screening model. Although the number of samples in our dataset, from a modelling prospective is small, the simplicity and the interpretability of the developed model are the positive aspects of the new model.

This part of the thesis provides a relatively comprehensive review upon WOE as inferred from the present large data and potentiality of metal oxide NMs chemical descriptors for assessment of DNA damage. It can be used as an informative platform in genotoxicity studies of metal oxide NMs. Such a combined approach can assist in providing useful insight about parameters that affect genotoxicity and thus provide guidance for the selection and/or design of safe NMs. In addition the identified quantum-mechanical descriptors in the classification model can be useful to prioritise the study of the NMs, which may lead to high risk for human health, especially in regulatory purposes.

CHAPTER 3

3. Materials and Methods

3.1 Part 1 – Carcinogenicity Models

3.1.1 Carcinogenesis Data Sources

3.1.1.1 ANTARES Carcinogenicity Dataset: Rat Carcinogenesis Learning Set

The first carcinogenicity model is developed on the basis of carcinogenicity database of EU-funded project ANTARES [107]. This data base contains rat carcinogenesis data (presence of carcinogenic effects in male or female rats). The ANTARES carcinogenesis data base is a collection of the EU-funded project CAESAR data set and the "FDA 2009 SAR Carcinogenicity—SAR Structures" data base. The

CAESAR toxicity values were originated from the Distributed Structure-Searchable Toxicity DSSTox database, which was built from the Lois Gold's Carcinogenic Potency Database [112]. The compounds with a definite TD50 (which is the dose that produces an incidence of 50% of the tumors in animals) value for rat in this dataset were labeled as carcinogenic, while the remaining were labeled as noncarcinogenic. 805 chemicals with carcinogenicity data were obtained from the CAESAR data set and 738 compounds are added from the "FDA 2009 SAR Carcinogenicity—SAR Structures" database using the Leadscope database [87]. A total number of 1543 compounds constituted the ANTARES dataset.

3.1.1.2 ISS Carcinogenicity Database and Carcinogenicity Genotoxicity Experience Dataset: Different Species Carcinogenesis Learning Set

For the learning set of the second prediction model we combined two carcinogenesis data sets. The ISS Carcinogenicity (ISSCAN) [109] database provided by the Istituto Superiore di Sanità is designed for the carcinogenicity predictive models. Most of the chemicals in the ISSCAN database are labelled as carcinogens by various regulatory agencies and scientific bodies. The database has been specifically designed as an expert decision support tool and contains information on chemicals tested with the long-term carcinogenicity bioassay on rodents (presence of carcinogenic effects in male or female rats and mice).

This carcinogenicity dataset contains 622 carcinogens, 210 noncarcinogens and 58 equivocals. We eliminated the chemicals with equivocal data, as we needed a definite carcinogenic effect for each data point. We merged the positive and the negative compounds with the ISSCAN database and the Carcinogenicity Genotoxicity eXperience (CGX) database. The CGX database was created by Kirkland et al. [108] and did not contain any equivocal result.

All compounds in the combined dataset have been checked for their consistency between the two sources. We found 651 compounds in common, 15 of them with inconsistent carcinogenicity values. These compounds have been removed from the combined dataset. In the present study, this combined dataset is conventionally called ISSCAN-CGX.

3.1.2 Comparison between the ANTARES Dataset and the ISSCAN/CGX Dataset

We compared the ISSCAN-CGX dataset with the ANTARES carcinogenicity dataset prepared for the development of the first model. The result of the check was 105 compounds with conflicting values. In order to develop a more conservative model, we decided to remove only 15 compounds with positive result in the ANTARES dataset and negative results in the second dataset, and left as carcinogenic those that had carcinogenicity result the opposite way. Consequently, there are 90 positive compounds in the ISSCAN-CGX database which are negative in the ANTARES dataset. Afterward, we checked and cleaned the structures manually, and by the help of the istMolBase [113] and InstantJChem [114] software formed the final dataset. In addition, we kept only the compounds with connected molecular structure; those which had unconnected structures have been removed from the dataset. The final dataset contained 986 compounds with 734 carcinogens and 252 noncarcinogens. For compound in the list these information are available: a chemical name, a CAS number, a SMILES [115], and its categorical designation (i.e., carcinogen or noncarcinogen).

3.1.3 Data for Model Validation

3.1.3.1 ECHA Database

In order to evaluate the two new carcinogenicity models developed, we prepared an external test set from carcinogenicity data in the eChemPortal inventory [116].

The constraints of the first query were: Study result type: experimental result; Reliability: 1 and 2 (1 = reliable without restrictions. 2 = reliable with restriction); Species: mouse and rat; Maximum number of studies: 4.

The second query consisted of: Study result type: experimental result; Reliability: 1 and 2; Species: mouse and rat; Sources: any guideline and exposure route.

The result of the first query was 308 compounds, whereas the second query returned 166 compounds, which were mostly in common with the chemical compounds of the first query. We manually examined the studies conducted for the first list of compounds, then we looked into the CLP inventory [117] for the positive chemicals collected by the previous queries. Inside the CLP inventory we found 68 compounds, which were already present in our data collection. This search confirmed the carcinogenic property of these compounds. The dataset consisted of 64 positive compounds, 169 negative compounds, and 90 equivocal compounds. The equivocal results are due to the presence of conflicting information in different sources or different studies in the same source. It should be noticed that for already classified compounds (no conflicting information), the level of uncertainty in the assignment is not homogeneous, because some of the compounds were classified on the basis of a single study (i.e., data present in one single source).

From the reliability point of view, in the data collected in our dataset, 49 positive compounds have positive carcinogenic effect in at least two sources. Fifty-seven negative compounds are noncarcinogenic in both lists, and they are not present in the list of compounds retrieved from the CLP inventory. Sixty four compounds are considered as noncarcinogens because of the presence of only one single study in the two lists.

3.1.4 Active Molecular Fragments Identification by SARpy

The SAR in Python (SARpy) program is a Python script based on the OpenBabel chemical library. SARpy creates classification models by identifying active and inactive molecular fragments by mining the chemical structures in a learning set. These extracted molecular substructures in the form of SMARTS [118] may be exactly similar to the already known SAs or newly developed SMARTS that are associated with a particular biological, pharmaceutical, or toxicological activity.

The generated molecular fragments are of arbitrary complexity, and the fragments candidate to become SAs are automatically selected on the basis of their prediction performance in a learning set.

SARpy takes a learning set of chemicals, where the molecule structures are represented as SMILES [115] notations, along with their experimental activity binary labels (e.g. toxic/nontoxic, mutagenic/nonmutagenic). This data mining tool generates every possible substructure in the set and finds correlations between a particular molecular substructure and the activity of the molecules that contain it. This is achieved in three phases:

(1) Fragmentation: this novel, recursive algorithm considers every combination of bond breakages working directly on the SMILES string. During this procedure the rings are not fragmented, they remain entire.

(2) Evaluation: the predictive ability of each extracted potential SA is examined on the training set.

(3) Rule set extraction: a reduced set of rules is extracted in the form:

'IF contains <SA> THEN <apply activity label>';

Where the SA is expressed as a SMARTS string, for use by human experts or chemical software. SMARTS notations are text representations of substructures that allow specification of wild card atoms and bonds, which can be identified to formulate substructure queries for a chemical database. Those rules can be used as a predictive model simply by calling a SMARTS matching program. For the matching phase,

SMILES and the SMARTS strings are translated into graphs and the two graphs are compared to each other [119].

This approach has been tested on the mutagenicity endpoint, showing marked prediction skills and, more interestingly, bringing to the surface much of the knowledge already collected in the literature as well as new evidence.

To each SA extracted from the learning set a statistical value is associated: training likelihood ratio. The molecular fragments identified by SARpy as active or inactive rules are compared to the molecules in the training set. Considering the experimental label of each molecule, there are two possibilities for each SA: i) the active SA is found in a positive compound, called "true

positive", or in the case of inactive SA, it is found in a negative compound, called "true negative", ii) the active SA is found in a negative compound, it is called "false positive", or the inactive SA is found in a positive compound, called "false negative". These indicators are used to calculate the accuracy of each SA in predicting the target activity label. In the case of active SAs, the likelihood ratio, which is a measure of precision intrinsic to the test (not depending on the prevalence of activity labels in the training set), is used as in the Formula 1:

$$Likelihood\ ratio = \left(\frac{TP}{FP}\right) * \left(\frac{negatives}{postives}\right) \qquad \textbf{(Formula 1)}$$

The training LR as a statistical measurement helps the user to evaluate the relevance of each SA in the training set used for building the model.

SARpy can be also used as a prediction tool for structure-based classification models. A list of SAs presented as SMARTS can be loaded into SARpy and a list of molecules SMILES strings can be inserted as test set. SARpy calculates the confusion matrix on the basis of the prediction results and provides accuracy, sensitivity and specificity of the classification model as output of the model evaluation.

3.1.5 Extracting Active Fragments

3.1.5.1 R (Rat) Model

To obtain a more comprehensive collection of potential carcinogenic fragments, five learning sets were randomly created from the ANTARES carcinogenicity dataset with 1543 compounds, preserving 80% for the learning set and 20% for the evaluation set. In other words, for each model a random set of 20% of chemicals in the learning set was removed, with the remaining 80% of the compounds a model was developed and the activity of the compounds left out was predicted with the same model. We combined the five models and put together the lists of the potential active fragments, removed the duplicates and eliminated the SAs with likelihood ratio lower than two. We opted for the likelihood ratio threshold of two in order to retain the SAs that are statistically more significant. A measure of each fragment's association with biological activity is determined by SARpy as "training likelihood ratio," and it is given along with the list of the potential fragments or the rule set in the output. The likelihood ratio can be taken into

account to determine the goodness of a SA identified by SARpy. Even if a SA that is associated with activity (i.e., carcinogenicity) is present in a molecular structure, the molecule may contain other fragments that make it inactive (i.e., noncarcinogen), thus the specific SA might not be expected to be found only in active compounds. This evidence is the basis of the determination of the likelihood ratio.

Using the SARpy software, each chemical in the learning set was fragmented *in silico* into all possible fragments meeting user-specified criteria. For this study we extracted only the "ACTIVE" fragments (or SAs) and the default values for the minimum and maximum number of atoms in a fragment were set for the fragment extractions of each model (minimum = 2; maximum = 18). Another configuration to establish by the user is the minimum number of compounds in the learning set in which an active (or inactive) fragment is found. In our analysis, the minimum number of compounds that contain a potential active fragment was set to three.

Conventionally, in this study we call this model R.

3.1.5.2 E (Expert) Model

SARpy was used for model development and statistical analysis using the ISSCAN-CGX dataset.

The extraction settings are as follows: the minimum number of atoms in a fragment is equal to four, whereas the maximum number of atoms is equal to 10, and the minimum number of compounds containing the active fragment is six. These configurations have been set in favor of a model with a more balanced sensitivity and specificity values. In order to assess the predicitivity of the model, statistical analysis have been conducted in terms of accuracy, sensitivity, and specificity using cross-validation routine as an internal evaluation, in addition to an external evaluation using an external test set. In this article, we name this model E.

3.1.6 Internal Evaluation of the Models

Accuracy, sensitivity, and specificity have been determined for the internal evaluation of each model using the SARpy program. For the internal validation, five-fold cross-validation routine was conducted for each model. In the five-fold cross-validation the learning set is randomly

partitioned into five equal sized subsets. For each iteration, a single subset of chemicals was retained as the validation data for testing the model, and the remaining subsets were used as training data. The cross-validation process was repeated five times (the folds). The evaluation results of five iterations were then averaged to produce a single estimation. Accuracy, sensitivity, and specificity of the internal evaluation are assessed in addition to the MCC.

3.1.7 External Evaluation of the Models

The predictability of the models has been evaluated on two external test sets: the first external set is the dataset used as the learning set of the opposite model (e.g., for the R model we used ISSCAN-CGX dataset and vice versa), and the second dataset is a collection of 258 compounds collected from the eChemPortal inventory. Accuracy, sensitivity, specificity, and the MCC for the external evaluation are determined using SARpy. Although the external evaluation is considered the best mean for the assessment of the predictive ability of a (Q)SAR model [120,121], the results of the external evaluation of any model are highly related to the relative similarity of the external evaluation set in relation to the learning set.

3.2 Part 2 – ToxDelta

3.2.1 Database of Active and Inactive Structural Alerts

Benfenati et al. [122] in a previous study have collected an expanded list of mutagenicity SA. This rule set is implemented in the ToxRead [123] software for the mutagenicity assessment within a read-across approach. The rules identified and collected in this collection are associated with Ames test bacterial mutagenicity and are based on more than 6000 chemicals from different chemical classes. These rules belong to both categories mutagenic and nonmutagenic and are sorted in a hierarchical way. ToxRead utilizes these rules in order to identify the active or inactive mutagenic substructures present in the target compounds. The hierarchical characteristic of the SAs in the list makes the process of rule search and identification more systematic. In other words, the exact SAs that match the target molecules are identifies and then more generic ones, which may match with the target molecule. In addition to the mutagenic active and inactive SAs, the rules present in this data set are accompanied with the exceptions and modulators of activity. From a toxicity prediction point of view, the identification of these SA in the compounds under assessment helps the expert to address the toxicity or nontoxicity of a molecule concerning the influence of each SA found in the molecular structure. Each SA is associated with its accuracy and p-value as statistical characterizations. The accuracy value indicates the precision of the SA as a potential agent in causing mutagenicity or decreasing the risk, considering the number of the molecules including this SA in the original training set. This set of SAs are implemented in the ToxRead program. There are more than 800 SAs present in this dataset with a high level of details such as accuracy and statistical significance.

3.2.2 The MCS Algorithm

While ToxRead analyses the similarities between a target molecule and the source molecules which are structurally similar, ToxDelta identifies the dissimilarities between these molecules. The identification of the dissimilar fragments is achievable after extraction of the MCS between two molecular structures. In our study, the degree of similarity between pairs of molecules is based on their molecular structure. Molecular structures can be encoded in several computer formats with topological information about the atoms and bonds of a molecule, as well as other

chemical information such as charges, aromaticity, etc. We opted for SMILES strings [115] as presentation of the input molecule in our program. The algorithm proposed by the fmcs_R package [124] extracts the MCS part between two molecule graphs using a novel backtracking algorithm by constructing a search tree of correspondences between nodes of the two graphs representing the two molecules. Each node in the tree presents a set of atoms correspondences of the respective molecule and the connected sub-graph we are looking for are in fact leaves, and the deepest leaves in the tree are the MCSs found.

An important issue to be noticed about the algorithm of the fmcs_R package is that it consists of a further characteristic, with respect to the other MCS extraction algorithms, and that is its "flexibility". In fact, this algorithm gives the possibility to the user to search not only all the exact MCSs, but also the flexible ones, in which the type of a limited number of atoms and bonds can be different in the two MCS extracted from two similar molecules. The result of searching a flexible MCS, probably, are pairs of different MCS, each belonging to one of the molecules under investigation. This option although makes the process of similarity finding more flexible and the number of the results is more elevated than the exact MCSs between two compounds, it is not in line with our objectives. Since our purpose of the MCS extraction is identifying the dissimilar fragments, for being more precise, we need to stay on the idea of the "exact" MCS extraction.

From a toxicity point of view, it is important to pay attention to the aromatic and aliphatic rings of the compounds.

The role of the aromatic and non-aromatic rings as structural properties of molecules have been highlighted in peer review resources. Among the important lists of mutagenic and carcinogenic SAs, aromatic and aliphatic rings play an important role. Different forms of rings are present in the mutagenicity and carcinogenicity rules [43]. The fmcs_R package breaks the rings whenever it is necessary in order to find the greatest part in common between two graphs. This leads to a significant loss of structural information and consequently the implication of the extracted MCS which is meant to be equal for both molecules may differ for each compound. Considering this important issue, we decided to add a new restriction to the fmcs_R algorithm. This restriction is to keep all the rings present in the target and the source molecule entire and do not partially

select a number of atoms of a ring. The algorithm selects the whole ring and add it to the MCS whenever this is possible, otherwise it does not take the ring into consideration for the MCS. The restriction of not breaking the rings in the process of MCS findings is added during the process of atom selection in the algorithm. Before adding any atom to the string of the MCS the algorithm checks if the atom belongs to a ring in both molecules and gathers all the information about the corresponding ring in the target and the source molecule. Only if the rings in both molecules have the same properties, (e.g. the number of atoms, the type of atoms, the type of bonds), the whole ring is added to the MCS. At this point we can extract the structural differences between the two compounds under investigation: we overlap each graph with the MCS and highlight all the sub-branches not in the MCS (Figure 6-page 59).

3.2.3 ToxDelta Implementation

ToxDelta is implemented as a complementary section to the ToxRead program. ToxRead associates the most similar molecules present in its data base with the target molecule, pointing out the mutagenic (or nonmutagenic) fragment(s) as toxicity rules present in both the target and the similar chemical compounds. ToxRead is a read-across tool based on similarity and identifies the mutagenic or nonmutagenic SAs in common between the target and the source chemicals. These SAs belong by definition to the MCS of the pair of compounds under investigation. Both tools operate by taking advantage of the list of identified mutagenic and nonmutagenic potential SAs. The user who wants to assess the mutagenicity effect of a molecule can evaluate the results obtained from ToxRead and the evidences gained from ToxDelta and make a decision regarding the mutagenicity of the target molecule, in a WOE approach. Figure 7 (page 60) shows the two phases of implementation of ToxDelta: i) the SA list creation and improvement, and ii) the evaluation of the mutagenicity of the target molecule considering its similarities and dissimilarities comparing to other known molecules with a high structural similarity. The evaluation of the degree of similarity in ToxDelta relies on the ToxRead program, which uses an ad hoc similarity algorithm described elsewhere [125]. A stand-alone version of ToxDelta is accessible on the VEGA home page (https://www.vegahub.eu/). Examining the results of ToxRead and ToxDelta will allow a thorough investigation of a compound with unknown mutagenicity property. As a first step, ToxRead investigates all the

SAs present in the target molecule and represents the most similar molecules to the target molecule which share the same SA with the target molecule. After this assessment, in order to have a further look, the user can select each of the source molecules and check the dissimilarities of the target and the source compounds. These tools help the researchers to identify the similar and dissimilar moieties. ToxDelta provides the most similar SAs for each of them in the collection of the known SAs for each moiety found. To obtain a conceivable result, the structure of the target and the source molecules in the comparison need to be sufficiently similar. If the structures of the molecules compared by ToxDelta do not share a significant MCS, the dissimilarities may not be interpretable to an acceptable level. In other words, whenever the structures of two molecules are strongly dissimilar, the user may not expect a significant MCS. In this regard the VEGA chemical similarity index [125] is used as a screening before applying the MCS approach.

In case the identified dissimilar fragment in the target of the source molecule is a SA, and it belongs to the list of SAs with an assigned accuracy and p-value information, there are three possible scenarios that can be associated with the dissimilar fragment:

1. The SA is an active substructure with strong potential to increase toxicity;
2. The SA is an inactive fragment with strong potential to decrease toxicity,
3. The SA has no relevant impact on the effect.

In case 1 and 2 there is more probability that the dissimilar fragment affect the whole molecule, while in case 3 the dissimilar substructure does not cause any toxicity or nontoxicity in the effect of the molecule. In all three cases the software will provide documentation about each SA found as a dissimilar fragment in the target molecule. From regulatory point of view, documentation is an important factor in toxicity assessment of the compounds and the acceptance of the read-across results. ToxDelta makes a thorough search in the list of identified SAs and provides not only the exact SA found as dissimilar fragment in the molecules but also the SAs which are so similar to the identified SA, in order to give more information to the expert for the assessment of the new molecule. This whole list of SAs is used by ToxDelta to assess whether the fragments resulting from the subtraction of the MCS from the molecule are associated with an increased or decreased or neutral effect.

The output of ToxDelta consists of all the possible MCSs extracted from two molecules of interest. The user can select one of these MCSs and evaluate the dissimilarities calculated based on the selected MCS. The different fragments present in both molecules, are the result of the subtraction of the MCS and the target or source molecules.

Figure 6. The MCS between two molecules is shown with bold lines, and the other branches are the differences

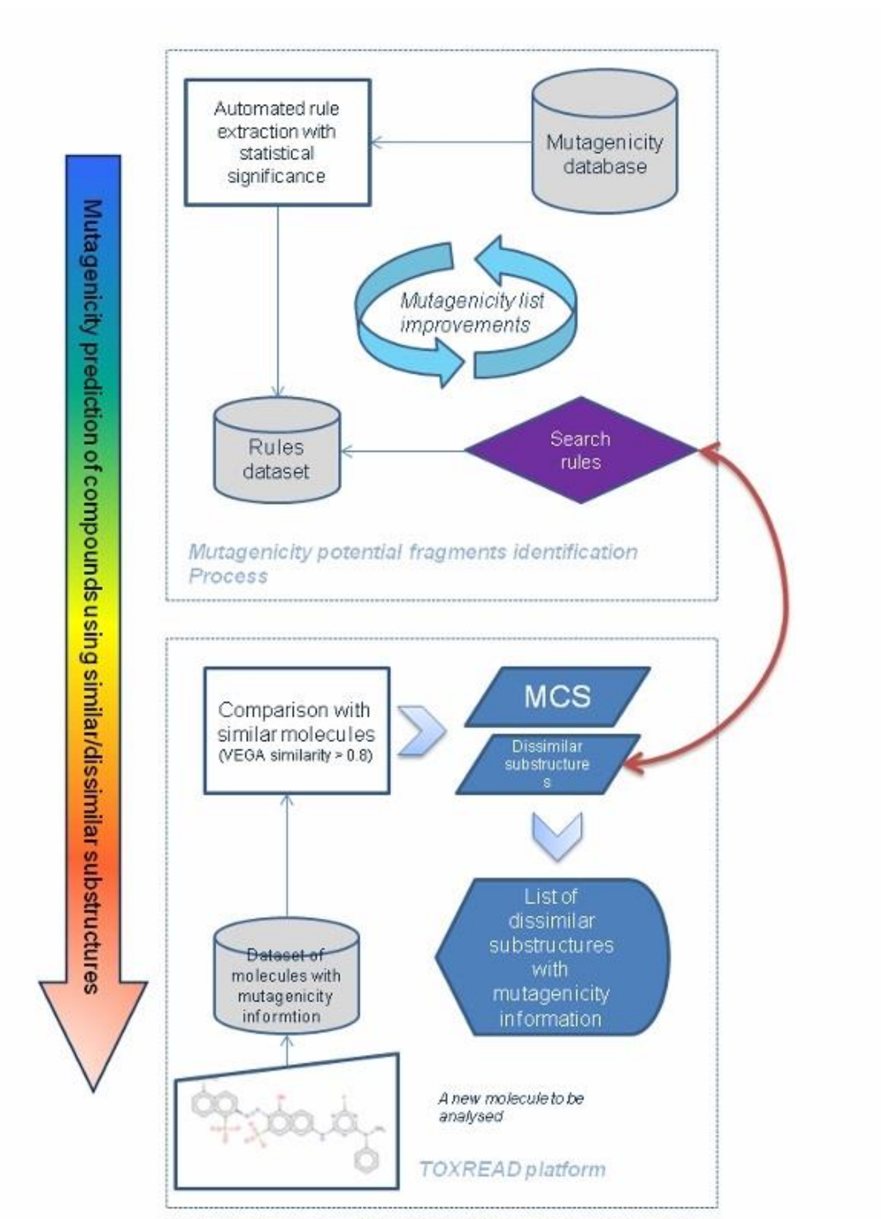


Figure 7. The flow chart of ToxDelta: the molecular similarity/dissimilarity structure analysis software for the mutagenicity endpoint

3.3 Part 3 – Metal Oxide NMs Genotoxicity Model

3.3.1 Data Collection and Assessment

Metal Oxide Nanomaterials and Genotoxicity Data

The list of metal oxide NMs considered in this study consists of: Al_2O_3, NiO, Co_3O_4, CuO, Fe_2O_3, Fe_3O_4, TiO_2, ZnO, SnO_2, V_2O_3, V_2O_5, MgO, ZrO_2, CeO_2, and Bi_2O_3. These metal oxide NMs with different chemical core compositions have been selected to study their genotoxicity effect towards humans. Although SiO_2 is technically a metalloid [126] it is considered a metal oxide NM in different nanotoxicity peer review documents [127]. Considering the similarities between silicon oxide and other metal oxides, we also considered SiO_2 as a metal oxide. Previously, our group conducted a study in which a collection of metal oxides with their Comet assay results were published [128]. These results are reported in the Appendices section. (Table S1.A and Table S1.B, Appendices)

3.3.1.1 Assessment of the Experimental Protocol

The different protocols and standards of the laboratories, result in heterogeneous results. Consistency and integrity are two challenging issues in this field. Consequently, the raw collected data is not suitable for building (Q)SAR models, since the data lines are now comparable [129]. Therefore, the experimental collected data were evaluated in order to put together data from different laboratories with the compatible protocols. We assessed 103 studies obtained from 75 publications for the assessment of metal oxide NMs in the *in vitro* Comet assay.

3.3.1.2 Prior Chemical Characterization

The assessment of the peer review documentation has been done on the basis of the criteria stabilized within the NanoPUZZLES project [130].

The same criteria in the NanoPUZZLES project has been applied to our data, and additional assessment criteria to NanoPUZZLES analysis were introduced in the current work.

The physico-chemical characterization of the NMs is reported in Table 4 (page 68) and should meet at least points 1, 2 and 3 of the following list:

1. Chemical composition and purity
2. Surface area
3. Particle size/size distribution
4. Crystal structure
5. Surface charge/zeta potential
6. Aggregation/Agglomeration status in the relevant medium. In principle, these are different phenomena (strong/weak binding between NMs), but are not always differentiated [131].

3.3.1.3 Minimum Criteria for the *In vitro* Comet Assay

We considered only the studies of the Comet assay or its variations as described by Singh et al. [132]. To each data point of 103 case study a list of queries have been applied. This list is reported below.

1. Whether the Comet assay was performed according to the guidelines presented by Singh et al. [132] with or without minor modifications.
2. Whether any of the following variations of the Comet assay were performed: with addition of 8-oxodguanine (8oxodG), formamidopyrimidine DNA glycosylase (FPG) and endonuclease III (endoIII). (It is to be noted that Comet assay can be performed by adding lesion specific bacterial glycosylase/endonuclease enzymes after analysis. These tests will detect a broader class of oxidative DNA damage bases. Whether the assays mentioned above were performed or not was verified).
3. Whether the pH of the electrophoresis was alkaline
4. Whether the concentrations of the tested NM were expressed in one of the three modalities as indicated in Table 4 (page 68).
5. Whether a cytotoxicity test was performed
6. Whether uptake into the cells was evaluated
7. Whether a dose - response analysis was performed
8. Whether positive and negative controls were used

9. Whether the Olive tail percentage was measured by analysis of at least two replicates of 50 cells
10. Whether there was an exposure of at least 3h

3.3.1.4 Classification of the Data Based on the Assessed Quality

We examined the experimental protocols of the Comet assays according to the questions listed in the "Minimum criteria for the *in vitro* Comet assay" section. The answers to these queries were "yes" or "no". The results of this assessment is classification of the data into three classes of reliability:

1st class data (high reliability): if at least questions 3 to 7 were answered "yes", the data line in classified into this class.

2nd class data (moderate reliability): Whenever cytotoxicity was not assessed (question 5), the corresponding meta data were classified into class 2 and less reliability value is assigned to the data of the second class in the final overall assessment of the genotoxicity of a specific metal oxide with the same core composition.

3rd class data (low reliability): data without any dose-response studies performed (NMs of only two different concentrations were tested) or DNA unwinding performed in non-alkaline pH (question 7 and 3).

3.3.2 Weight of Evidence Approach in the Evaluation of the Data Set

To conclude and overall genotoxicity property for each metal oxide NM under investigation we employed a quantitative WOE approach. It is demonstrated that the quantitative WOE approach is the best strategy where individual lines of evidence are integrated by an authoritative expert to form a conclusion [133].

3.3.2.1 Assessing Data for the Same Core Composition but Different Size Range or Crystallography Reported in One Single Publication

In our study, we assigned an overall positive sign to a NM with the same core composition and size, were at least one positive result in a study was reported for the concerned metal oxide NM. Whereas, a negative sign was assigned to a NM, whenever all the studies reported a negative outcome for that specific metal oxide NM.

3.3.2.2 Assessing Data Derived from Different Publications for the nanomaterials of the Same Core Composition

The data points with "highly reliable" property had more priority in the assessment of the data gathered from the publications for the NMs. Whenever the results were conflicting the judgement was added to by second and third class data. In addition, where the characterization and the reliability class of the Comet assay were equal, we considered the results of the majority of the data lines. We eliminated one of the data points from the final overall dataset, as minimum criteria for the assay performed could not be assigned to any of the classes of reliability (extracted from Sekar et al.[134]). The numbers of genotoxic and nongenotoxic results for each metal oxide of the same core composition are presented in Table 5 (page 69). Each single datum point reported here is considered as one genotoxic or nongenotoxic call.

3.3.3 Case Studies for Illustrating the Weight of Evidence Evaluation

We selected three case studies in order to illustrate the WOE approach applied in the present study. The final judgement for each metal oxide NMs, of the same core composition, is reported as the "overall assessment" in Table 5 (page 69).

Case study 1

The TiO2 NPs have been tested in the Comet assay more often than other metal oxides (37 reports among which 31 reports were genotoxic and six were nongenotoxic). Hence, our overall assessment of the TiO2 NM, according to expert judgement, is that it is genotoxic. There were 16 highly reliable data points according to the criteria defined in Section "Minimum criteria for the *in vitro* Comet assay"; genotoxicity was seen in 12 out of these 16 data points.

Case study 2

The Al2O3 NPs were designated as being genotoxic. However, there are two Comet assay studies with nongenotoxic results for aluminium oxide: Kim et al.[135] and Demir et al.[136]. The extent of characterisation and minimum Comet protocol were not completely met by the study reported by Demir et al.[136] (e.g. it lacked a dose-response relationship study); thus, greater reliance is placed on the highly reliable results presented by Kim et al.[135].

Case study 3

Overall six reports for Fe_2O_3 NMs were available. Among these publications the results published by Auffan et al.[137] fulfilled all the criteria to be highly reliable data. This study was the only one to use the positive and negative controls as well as fulfilling the rest of the requirements as described in Section "Minimum criteria for the *in vitro* Comet assay". Guichard et al.[138] provided details of the assay performed that satisfy the criteria for highly reliable meta data and agree with the assessment of nongenotoxicity. Of all the six studies, only these two provided results of uptake studies. Furthermore, in support of our conclusion that Fe_2O_3 is nongenotoxic, the same result was found in most the reports of data with moderate reliability. The same reasoning approach was used for Fe_3O_4, ZnO, SiO_2 (Table 5-page 69). Data for the Fe_3O_4 NPs were reported in eight studies. According to Table S1.A (Appendices), data extracted from Guichard et al.[138] and Könczöl et al.[139] were highly reliable. Nevertheless, the two studies reported contradictory results. Other five data points fired the NM to be genotoxic. These data points met all the criteria to be highly reliable except for uptake studies. Expert WOE results for Fe_3O_4 were genotoxic. For ZnO NPs highly reliable data points assigned genotoxic call [140].

The same approach has been undertaken for the reports that included NiO, Co_3O_4, CuO, V_2O_3, V_2O_5, MgO, ZrO_2, Bi_2O_3, and SnO_2 as shown in Table S1.A (Appendices). If there were no highly reliable data points available, moderately reliable data were taken into consideration.

3.3.4 Computational Analysis of nanomaterials Structure and Descriptor Generation: Quantum-chemical Descriptors

The computational analysis of the structure of the metal oxides and the calculation of various molecular descriptors have been successfully conducted in the present study. A total of eleven descriptors were calculated by quantum chemical methods (quantum-chemical descriptors) for all the metal oxide NMs present in our data set.

Quantum-mechanical calculations of metal oxide clusters were performed based on experimental crystal lattice parameters obtained from the Crystallographic Open Database [141]. We adapted the method presented by Gajewicz et al.[142]. In order to have cubic clusters with an acceptable size that represent the molecular models of the NMs studied, the lattice parameters

were then increased in all three dimensions. We used GaussView for the generation of the molecular structures of all examined metal oxide clusters [143].

The quantum-mechanical calculations of the descriptors has been performed on the clusters of metal oxides in two phases: i) optimization of the molecular geometry, and ii) calculation of quantum chemical descriptors based on the optimized geometry.

It is important to notice that the generated clustered are too large for analysis by ab initio methods. For this reason, the semi-empirical level of theory has been employed, utilizing the efficient PM7 that has been re-parameterized for the elements considered in this investigation [144]. Additionally, the PM7 approach outcomes are more accurate compared to the DFT level, since it uses a novel parameterisation of the previously used PM3 Hamiltonian [145]. The PM7 method is implemented in the MOPAC2012 software package [144].

The following molecular descriptors were calculated for each metal oxide NM: heat of formation (HF), dipole moment (μ), total energy (ET), electronic energy (EE), the total solvent accessible surface area of the cluster (SASA), energy of the highest occupied molecular orbital of the oxide cluster (EHOMO), energy of the lowest unoccupied molecular orbital of the oxide cluster (ELUMO), ΔH of cluster and molecular weight of metal oxide cluster (MW). The total solvent accessible surface area (SASA) of the cluster was calculated by COnductor-like Screening MOdel, implemented in MOPAC 2012. The definition of each descriptor and the results are reported respectively in Table 6 (page 70) and (Table S3, Appendices).

3.3.5 Classification SAR Modelling Methods

The limitation in the number of data samples is an unavoidable problem with the metal oxide NM databases. The restriction of size is a challenge in building a reliable model of genotoxicity with high prediction accuracy. Even if a high number of molecular descriptors is calculated for the small data set of NMs, still we deal with the issue of "under sampling induced collinearity", which means a high degree of collinearity in descriptors [146,147]. Collinearity will be present in the model as the number of samples is very small compared to the number of descriptors. Additionally, other problems such as over-fitting and noise in the data with negative effects on the model will arise. Considering the abovementioned complications, in order to find the most

appropriate model to fit the data, it is better to focus on a limited set of hypotheses. In other words, in case of small data, it is better to start from a small set of possible hypotheses, e.g. a set of decision trees with depth <= four. Thus, we opted for a simple tree classification analysis for (Q)SAR modelling of our data set, in particular, Recursive Partitioning and Regression Trees (rpart) model was used to classify the data set.

3.3.5.1 Recursive Partitioning and Regression Trees

The rpart programs build classification and regression models in two phases and the result is a binary tree. To build the tree classification model the first phase is identifying the variable which contributes the most to the splitting the data into two groups. After dividing the data into two groups, the algorithms continues the splitting separately for each group. The procedure continues recursively until each group contains a minimum number of samples or no more improvement can be achieved. During the second phase, a cross-validation evaluation is performed on the data to trim the full tree and make is simpler [148].

Considering the small data set of metal oxide NMs with their associated set of their quantum-mechanical descriptors and the classification endpoint we need to model, the factor of “randomness” is likely to play a role in the built model. To overcome this situation, we decided to develop a model to analyse the importance of each variable in relationship with the genotoxicity property of the NMs, rather than a model to estimate the genotoxicity of the metal oxide NPs. The (Q)SAR models in addition to their predictive ability, help us to identify the more effective physico-chemical attributes of a chemical related to toxicological and biological properties of the substances. In the present study, (Q)SAR models are employed to study the effect of each quantum-chemical descriptors in amplifying or reducing the genotoxicity of the NMs. Considering the limitations mentioned above, we decided to use all the data as training set and study the importance of each descriptor in amplifying the genotoxicity property of the metal oxide NPs. All the quantum-chemical descriptors have been standardized in the data set prior to the modelling process. All analyses were done in R version 3.2.3 (R Foundation for Statistical Computing, Vienna, Austria), using the ‘rpart’ library.

Table 4. Criteria for the usefulness and quality assessment of the data set for the (Q)SAR modelling: extent of Comet assay conditions checklist. General parameters have been used to assess each data point and the results are reported in Table S1.A (Appendices) where all questions are answered in a yes or no fashion.

General parameters	Further details to assess
Comet protocol type:	I) The pH of unwinding: alkaline, neutral, very alkaline. II) Incubation with the enzymes: FPG, 8oxodG, Endo III.
Concentrations expressed in at least one of the units:	I) Mass per volume, per area, per cell (µg/ml, µg/cm2, µg/cell) II) Number of NMs per ml, per cm2, per cell (ENMs/ml or ENMs/cm2 or ENMs/cell) III) Surface area per ml, per cm2, per cell (cm2/ml or cm2/cm2 or cm2/cell)
Cytotoxicity tests performed?	
Performed trend test for dose-response relationship?	
Microscopic analysis in the Comet assay: Analyzed at least 50 Comets per gel divided on two different slides (parallel gels per sample)? Comet count performed at least by one of the methods?):	I) % DNA in the tail II) Tail length III) Tail moment IV) Tail intensity (classified as belonging to one of five classes depending on their tail intensity?)
At least 3 hours for treatment time was respected?	
Performed comparison between treated samples and controls?	I) Positive control II) Negative control III) Both negative and positive controls
Information on uptake (demonstrated cellular uptake?)	

Table 5. Comet assay experimental results for all selected metal oxide nanomaterials used for (Q)SAR modelling*.

No	Metal oxide	Number of genotoxic reports	Number of non-genotoxic reports	Overall assessment**
1	Al_2O_3	1	1	+
2	NiO	1		+
3	Co_3O_4	2		+
4	CuO	6	2	+
5	Fe_2O_3	1	5	-
6	Fe_3O_4	6	3	+
7	TiO_2	32	6	+
8	ZnO	16	1	+
9	SiO_2	3	9	-
10	V_2O_3	1		+
11	V_2O_5		1	-
12	MgO		1	-
13	ZrO_2		1	-
14	CeO_2	5	1	+
15	Bi_2O_3	1		+
16	SnO_2		1	-

* Data were extracted from [128].

** The "positive" and "negative" signs are assigned according to the number of genotoxic and nongenotoxic "reports" per each NM. The assessment column represents the variable used to model, based upon the global evaluation (weight of evidence) of all the reports related to a single NM (i.e. row): "+" means positive, i.e. genotoxic, whereas "-"means negative, i.e. not genotoxic.

Table 6. Acronyms, short definitions and units of the molecular descriptors calculated by MOPAC2012.

Symbol	Descriptors	Unit
H_F	Heat of formation	Kcal/mol
TE	Total energy of the oxide cluster	Ev
EE	Electronic energy of the oxide cluster	Ev
Core	Core-core repulsion energy of the oxide cluster	Ev
σCOSMO	Surface charge distribution based on Conductor-like Screening Model	Cubic Angstroms
COSMO-SA	Area of the oxide cluster calculated based on COSMO	Square Angstroms
IP	Ionization Potential	Ev
HOMO	Energy of the highest occupier molecular orbital of the oxide cluster	Ev
LUMO	Energy of the lowest unoccupied molecular orbital of the oxide cluster	Ev
No.Fl	Number of Filled Levels	adimensional
MW	Molecular Weight	g/mol

3.4 Weight of Evidence Approach in the Analysis of Results of Different In Silico Methods for the Mutagenicity Assessment of Chemicals

In a WOE approach the prediction results of two (Q)SAR platforms which include nine mutagenicity models are assessed to reach a conclusion about the mutagenic effect of two chemical substances as case studies. The results of each mutagenicity model are considered pieces of evidence. The goal is to integrate the results to reach a conclusion on mutagenicity of the chemical under investigation. These pieces of evidence form a line of evidence to be used in further investigations together with other types of lines of evidence such as *in vitro* or *in vivo* mutagenicity results to help the assessors to reach a reliable answer for a toxicity question. Two drugs are selected for the present practice: Valproic acid and Diclofenac.

Two methodologies are integrated for the proposed WOE framework. The results of the (Q)SAR and read-across *in silico* models are documented and integrated and additionally, the most similar compounds identified by the (Q)SAR platforms and ToxRead are analysed. The comparison between the target chemical and each individual source chemical is conducted by means of ToxRead and ToxDelta. The identified SAs as dissimilar fragments in the structure of each chemical are studied to explore their role in affecting or reducing toxicity (in this case mutagenicity effect).

CHAPTER 4

4. Results and Discussions

4.1 Part 1- Carcinogenicity Models

4.1.1 R Model

From each training set a collection of active SAs has been extracted. These collections or active rules of molecular substructures are the new models. Each SA is associated with a likelihood ratio, which is a statistical value for illustration of the goodness of the rule. The final model which is a result of merging all the rule sets consisted of 127 active SAs. Table 7 (page 75) shows the statistics of the prediction results of five models developed base on five different splits of the ANTARES database. The performance of each model is evaluated using its own test set. The average of the predictive values of all the five models have been reported in Table 7 (page 75), as well. The averages of accuracy (Formula 2), sensitivity (Formula 3) and specificity (Formula 4) for the 778 compound internal cross-validation using five rule sets extracted from the ANTARES dataset were 71%, 73% and 69%, respectively. The average of accuracy, sensitivity and specificity for 337 compounds in the test set as an external validation of these models, were 63%, 63% and 62%, respectively.

The results of cross-validation of the R model on the whole training set were 66% accuracy, 83% sensitivity, 48% specificity and 0.34 the MCC (Formula 5) (Table 8-page 76). The R model produced better results for the external evaluation of the model using the ECHA database. In fact, analysis of the external validation for the R model demonstrated that the concordance between experimental and predicted value on the ECHA dataset is higher than using the ISSCAN-CGX dataset. The accuracy of the R model on the ECHA dataset was 67%, compared to 58% of accuracy for the ISSCAN-CGX dataset. The complete list of these alerts are presented in the VEGA platform.

4.1.2 E Model

SARpy extracted 43 active SA from the ISSCAN-CGX training set. Analysis of the cross-validation for the E model demonstrated that the second model produced an accuracy of 73%,

with a sensitivity of 77% and a specificity of 62% (Table 8-page 76). The MCC value for this analysis is 0.36. The accuracy values for the external evaluation of the E model on the ANTARES dataset and the ECHA database were 59% and 64%, respectively. Analysis of the external validations for the E model demonstrated that the model produced a higher sensitivity (77%) compared with the specificity (41%) of the R model. On the contrary, the specificity of the external evaluation on the chemicals from the ECHA database was higher (72%) compared to its sensitivity (48%) (Table 8-page 76). The complete list of the SAs present in this model is accessible through VEGA.

4.1.3 Analysis of the Combination of the Prediction Results of the R and E Models

In addition to the separate analysis of the prediction results of the R and E model, another combined evaluation has been conducted. In this analysis of the prediction results of the R model and the E model, we considered the final results as correctly predicted only in case both models have predicted them consistently. Table 9 (page 76) summarizes the results of combining the R and E model external validation predictions on the chemicals from the ECHA database.

The results showed that in case both models had a concordant result on a negative prediction the reliability of the results is higher than in case the positive predictions.

We observe an improvement of the results compared to the use of the individual models, for accuracy (72%) and specificity (79%). In fact, combining the predictions of the two models the MCC is increased to 0.37, compared to 0.31 for the R model and 0.20 for the E model. Only sensitivity is higher using the R model (62%). Thus, users may choose a solution or another depending if they prefer a conservative or a realistic assessment.

$$accuracy = \frac{TP+TN}{total\ number\ of\ predictions} \quad \textbf{(Formula 2)}$$

$$sensitivity = \frac{TP}{TP+FN} \quad \textbf{(Formula 3)}$$

$$specificity = \frac{TN}{TN+FP} \quad \textbf{(Formula 4)}$$

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}}$$ **(Formula 5)**

4.1.4 Fragments Analysis

4.1.4.1 Comparison of the SAs in the R and E Models

We compared all the SAs in the R and E models in order to identify the SA in common between the two models. The common SAs have been categorized into chemical classes. The SAs in the R model are presented with their ID number and written in order of their correspondence to the identical SAs in the E model.

1) Aromatic amine (R model: 6, 41, 36, 22, 10 / E model: 27, 31, 33, 38, 104)
2) Aromatic heterocyclic (R model: 12, 19, 2 / E model: 75, 108, 117)
3) Hydrazide (R model: 28, 27 / E model: 2, 50)
4) N-Nitroso (R model: 1 / E model: 8)
5) Phenyl-Hydrazine (R model: 32 / E model: 48)
6) α,β- Haloalkanes (R model: 25 / E model: 56)
7) Sulfite (R model: 8 / E model: 68)
8) Nitrogen Mustard like (R model: 11 / E model: 73)
9) Phosphonite (R model: 15 / E model: 98)

4.1.4.2 Categorization of the SAs in the R and E Models

All the SAs of the R and E models are categorized into chemical classes. The substructures within each category are presented with their ID number in their original rule set and are as follows:

Nitrogen containing substructures (Azo type):

1) Aromatic amine (R model: 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36, 37, 38, 40, 42, 83, 104, 110, 113 / E model: 6, 10, 22, 31, 35, 36, 41, 42)
2) Aromatic heterocycles containing Nitrogen (R model: 74, 75, 80, 81, 83, 95, 113, 122 / E model:12, 17, 43)
3) Azine (Hydrazine) (R model:46, 47, 49, 50, 51, 53, 54, 55, 101 / E model: 27, 32)
4) Azide (Hydrazide) (R model: 2, 3, 44, 45, 52 / E model: 3, 28)
5) Nitrosamine (R model:4, 5, 7, 9, 10 / E model: not found (NF))
6) Nitrogen or sulfur mustard (R model: 72, 73, 115 / E model: 11, 34)
7) Aromatic methylamine (R model: 30, 34, 36 / E model: NF)
8) Aliphatic N-Nitroso (R model: 62, 63/ E model: NF)
9) Aromatic Nitro (R model: 90, 123 / E model: NF)
10) 1 aryl 2 monoalkyl hydrazine (R model: 48 / E model: NF)
11) Aziridine (R model:120 / E model: NF)
12) Aromatic hydroxylamine (R model: 32 / E model: NF)

13) Diazo (R model:92 / E model: NF)
14) Aromatic Azo (R model: 71 / E model: NF)
15) Aromatic Nitroso (R and E models: NF)
16) Other substructures:
17) (1,2, and 3 membered) Aromatic Heterocycles (R model: 74, 75, 80, 81, 83, 90, 95, 103, 108, 113, 117, 121, 122, 123 / E model: 2, 12, 17, 19, 43)
18) Aliphatic halide (R model: 57, 58, 59, 70, 125 / E model: 18, 25)
19) Heterocyclic Alkane (R model: 84, 105, 109, 120 / E model: 23)
20) Polycyclic aromatic systems (R model: 39, 43, 60, 61 / E model: 30)
21) Sulfonate bonded carbon (R model: 67, 68 / E model: 8)
22) Epoxide (R model: 105 / E model: 23)
23) B propiolactone (R model: 114 / E model: NF)

SARpy had successfully identified most of the already known carcinogenic substructures that were presented by Kazius et al. In addition a number of SAs have been extracted by SARpy for the first time. Table 10 (page 77) demonstrates the new identified SAs that have been classified into seven chemical classes. The substructures within each category are listed with their ID number and are as follows:

1) Nitrosurea (R model: 12, 13, 14, 19 / E model: NF)
2) Nitrogen or sulfur mustard like (R model: 72, 115 / E model: 34)
3) Benzodioxole and Benzendiol (R model: 17, 18 / E model: 9)
4) Teritiary amine substituted by a Sulfur atom (E model: 24)
5) α,β-oxy and carboxy substitutions (R model: 20, 21, 76 / E model: NF)
6) α,β-haloalkanes (R model: 56, 69 / E model: 25)
7) Oximes (R model: 78 / E model: NF)

As an example, we illustrated the chemicals from which the SA 24 (from the chemical class teritiary amine substituted by a Sulfur atom) in the E model has been extracted (Table 11-page 78). It is important to notice that all the chemicals that contain the above mentioned SA in the ISSCAN-CGX data set are carcinogenic.

Table 7. R model internal and external validation for five different splits and the average of the model performance

		1° split (59 active rules)	2° split (65 active rules)	3° split (61 active rules)	4° split (58 active rules)	5° split (57 active rules)	Average
Learning set (778 compounds)	Accuracy	71 %	72 %	71 %	70 %	71 %	71 %
	Sensitivity	75 %	75 %	71 %	73 %	70 %	73 %
	Specificity	65 %	69 %	71 %	66 %	72 %	69 %
Test set (337 compounds)	Accuracy	63 %	60 %	64 %	65 %	62 %	63 %
	Sensitivity	68 %	58 %	62 %	67 %	61 %	63 %
	Specificity	56 %	63 %	66%	61 %	64 %	62 %

Table 8. R model and E model internal and external validation

	R model (127 active rules)			E model (43 active rules)		
	Cross-validation	external validation on ISSCAN and CGX data	external validation on ECHA data	Cross-validation	external validation on ANTARES data	external validation on ECHA data
Accuracy	66%	58%	67%	73%	59%	64%
Sensitivity	83%	76%	62%	77%	77%	48%
Specificity	48%	40%	70%	62%	41%	72%
TP[a]	651/783	593/735	55/89	562/735	599/783	43/89
TN[b]	367/760	142/254	119/169	157/254	315/760	121/169
FP[c]	393/760	112/254	50/169	95/254	445/760	48/169
FN[d]	132/783	142/735	34/89	172/735	184/738	46/89
MCC[e]	0.34	0.35	0.31	0.36	0.19	0.20

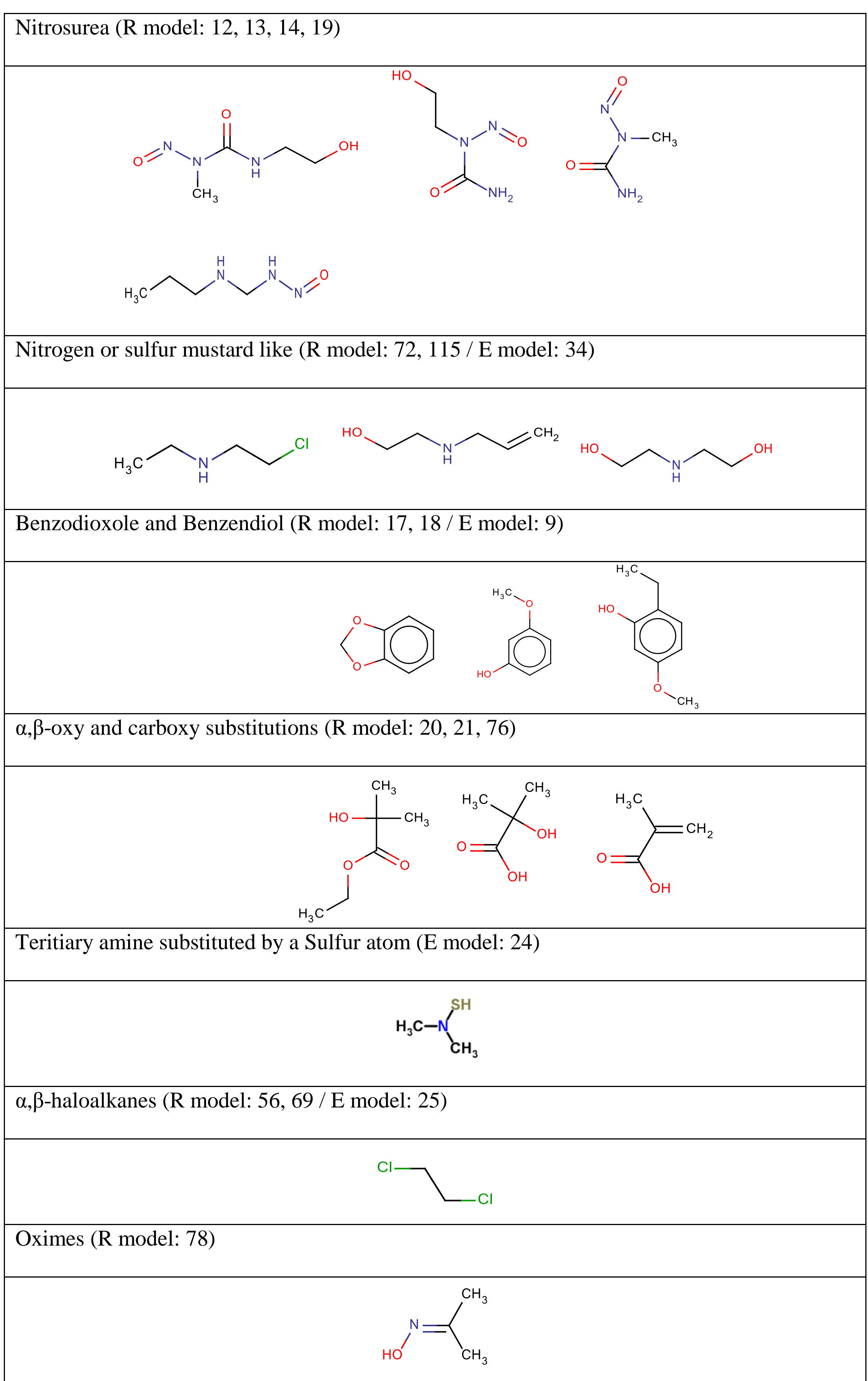
[a] True positive; [b] True negative; [c] False positive; [d] False negative; [e] Matthews Correlation Coefficient

Table 9. The combination of the predictions of the R and E models on the ECHA external validation set

Combined model	
TP[a]	33/89
TN[b]	96/169
FP[c]	25/169
FN[d]	24/89
Accuracy	72%
Sensitivity	58%
Specificity	79%
MCC[e]	0.37
Coverage	178/258

[a] True positive; [b] True negative; [c] False positive; [d] False negative; [e] Matthews correlation coefficient

Table 10. New carcinogenic structural alerts identified by SARpy in the R and E models



Nitrosurea (R model: 12, 13, 14, 19)
Nitrogen or sulfur mustard like (R model: 72, 115 / E model: 34)
Benzodioxole and Benzendiol (R model: 17, 18 / E model: 9)
α,β-oxy and carboxy substitutions (R model: 20, 21, 76)
Teritiary amine substituted by a Sulfur atom (E model: 24)
α,β-haloalkanes (R model: 56, 69 / E model: 25)
Oximes (R model: 78)

Table 11. Chemicals structures in the ISSCAN-CGX data set from which structural alert 24 has been extracted

O=C1c2ccccc2C(=O)N1SC(Cl)(Cl)Cl	O=C1N(C(=O)C2CC=CCC12)SC(C(Cl)Cl)(Cl)Cl	O=C1N(C(=O)C2CC=CCC12)SC(Cl)(Cl)Cl
O1CCN(C(=S)SN2CCOCC2)CC1	O=C(O)c1ccc(cc1)S(=O)(=O)N(CCC)CCC	O=C(O)c1cc(ccc1Cl)S(=O)(=O)N1CC(C)CC(C)C1

4.2 Part 2- ToxDelta

In the stand-alone version of ToxDelta, the user can insert two chemical compounds and compare their molecular structure. The two chemical compounds can be introduced as SMILES [115]. The MCS is the biggest common part between two molecules in input and is shown in Table 12 (page 85). Usually the assessment of a target molecule in a read-across approach is based on the comparison of the structure of the compound under investigation and another source compound with known toxicity property. In our program, the MCS is the common part in both molecules which is analysed and processed by ToxRead in the initial phase of read-across process. Indeed, the application of ToxDelta is useful for substances that are structurally similar. The MCS typically is an important part in read-across procedure. In this scheme, ToxDelta complements the conceptual strategy of ToxRead. The problem with the existing read-across tools is the risk of missing the differences between two similar molecules. The similarity should not minimize the fact that the possible opposed behaviour of the two similar compounds. In order to avoid the lack of attention to the opposite behaviour, ToxDelta makes a thorough assessment about the differences of the similar compounds under read-across examination. The theoretical basis is closely related to the SA paradigm. Thus, ToxDelta is an interdependent part of the ToxRead software, which exploits all the SAs of the target compound.

ToxDelta makes is possible to take a closer look at the two substances (i.e. the target and the reference compounds), in particular, when they may have opposite toxicological properties. Indeed, it should be reminded that ToxRead predicts the toxicological property of the target compound, and thus the predicted value of the target compound may be the completely different from the experimental value of the similar compound.

To investigate the utility of ToxDelta in read-across we provide two examples of mutagenicity endpoint. Each example includes two chemicals. For convention, we label one of the chemicals "Target molecule" and the other chemical "Source molecule".

4.2.1 Case Study 1: Benzodiazepine Derivatives

Target molecule 1: Diazepam

Systematic name: 1-methyl-5-phenyl-7-chloro-1,3-dihydro-2H-1,4-benzodiazepin-2-one

SMILES: O=C1N(c3ccc(cc3(C(=NC1)c2ccccc2))Cl)C

Experimental activity: nonmutagenic in Ames test [52]

CAS number: 439-14-5

Source molecule 1: Flunitrazepam

Systematic name: 1,3-dihydro-5-(o-fluorophenyl)-1-methyl-7-nitro-2H-1,4-benzodiazepin-2-one

SMILES: c12C(=NCC(=O)N(c1ccc(c2)[N+](=O)[O-])C)c1c(cccc1)F

Experimental activity: mutagenic in Ames test [52]

CAS number: 1622-62-4

The first pair are **Diazepam** (we suppose with unknown mutagenic property) as the target molecule and **Flunitrazepam** (mutagenic in Ames test) as the source molecule. We suppose that we have no information about the mutagenic effect of the target molecule and the aim is to investigate the possibility of assigning the mutagenicity property of the source molecule to the target molecule. The similarity index between the first pair of molecules is 0.871, and the MCS between the target and source molecule is extracted by ToxDelta ("1-methyl-5-phenyl-2,3-dihydro-1H-1,4-benzodiazepin-2-one") (Table 12-page 85). Since ToxDelta is a complementary tool implemented in ToxRead, prior to using ToxDelta the user have the possibility to illustrate the common SAs between the target and the source molecular structures. Indeed, ToxRead identifies the SA in common between the two molecules as following:

Name: SM203

Description: Sarpy alert n. 203 for NON-Mutagenicity, defined by the SMARTS: N(C)(CCN)c1ccccc1

Experimental accuracy: 0.52

Fisher test p-value: 0.55052

Since the SA in common between the two molecules is not an active mutagenicity rule, it is important to investigate the dissimilarities between the molecules. At this point ToxDelta is used to identify the differences and to extract all the dissimilar fragments, which do not belong to the MCS in both molecules.

ToxDelta extracts the following SA as one of the dissimilar fragments from **Flunitrazepam** from the mutagenicity SA dataset of ToxRead/ToxDelta [122,123]:

Name: SA27

Description: Nitro aromatic (Benigni/Bossa structural alert no. 27)

Experimental accuracy: 0.87

Fisher test p-value: < 10e-6

Given that there are no other significant SA in **Flunitrazepam,** the Nitro aromatic mutagenicity rule with 0.87 experimental accuracy is likely to be the reason of the mutagenicity effect of the source compound.
The other dissimilar fragment identified in the source molecule (Fluorobenzene) is also present in the target molecule (Chlorobenzene). This dissimilar fragment is halogenated benzene with a low experimental accuracy for mutagenicity effect, thus is not likely to trigger mutagenicity.

Name: SA31a

Description: Halogenated benzene (Benigni/Bossa structural alert no. 31a)

Experimental accuracy: 0.48

Fisher test p-value: 0.00735

Considering that no other significant mutagenicity SA is identified in the structure of **Diazepam**, the user may conclude that although the similarity index between **Diazepam** and **Flunitrazepam** is high (0.871), the molecular structure investigation of the two molecules do not provide evidence to the possibility of assigning the property of the source molecule to the target. Not being available any other SA with high mutagenicity accuracy inside the MCS or the dissimilar fragments extracted from the target after the comparison with the source, the user may conclude that **Diazepam** is not mutagenic. Indeed, we already have the Ames test value for **Diazepam**, which is negative for this endpoint.

As a conclusion, ToxDelta immediately reports as a key difference the presence of the nitroaromatic fragment, which is at the basis of the different mutagenicity value of the two substances.

4.2.2 Case Study 2: Androstane Derivatives

Target molecule 2: **Mepitiostane**

Systematic name: 5-alpha-Androstane, 2-alpha,3-alpha-epithio-17-beta-(1-methoxycyclopentyloxy)-

SMILES: O(C)C6(OC2CCC3C4CCC1CC5C(CC1(C)C4(CCC23(C)))S5)(CCCC6)

CAS number: 21362-69-6

Experimental activity: nonmutagenic in Ames test [149]

Source molecule 2: **Cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-**

Systematic name: Cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-

SMILES:

O=S5(=O)(CCS(=O)(=O)C35(CC1C4CCC(C(C)CCCC(C)C)C4(C)(CCC1C2(C)(CCC(CC23)Br)))

CAS number: 133331-34-7

Experimental activity: mutagenic in Ames test [149]

The second pair are **Mepitiostane** (we suppose with unknown mutagenic property) as the target molecule and **Cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-** (mutagenic in Ames test) as the source molecule. We suppose that we have no information about the mutagenic effect of the target molecule and the aim is to investigate the possibility of assigning the mutagenicity property of the source molecule to the target molecule. The similarity index between the second pair of molecules is 0.774, and the MCS between the target and source molecule is extracted by ToxDelta (Table 12-page 85). ToxDelta can be used as a stand-alone tool or an auxiliary tool inside ToxRead. ToxRead illustrates the common SAs between the target and the source molecular structures. Indeed, SAs in common between the two molecules identified by ToxRead are as following:

Name: SM153

Description: Sarpy alert n. 153 for NON-Mutagenicity, defined by the SMARTS: SCCCC

Experimental accuracy: 0.83

Fisher test p-value: 0.00005

It is evident that the common substructure between the molecules under investigation is not a potential mutagenic SA, thus the dissimilarities of the two molecules need to be explored for the potential mutagenic substructure.

ToxDelta identifies the androstane tetracyclic system as MCS shared by these two chemicals and extracts five fragments of dissimilarity (Table 12-page 85). Three of these are aliphatic rings: the thiirane, 1,1-dimethoxycyclopentane, and 1,3-Dithiolane 1,1,3,3-tetraoxide rings and two are aliphatic chains: the 2-methylheptyl group and a bromine atom, both linked to an aliphatic carbon ring. The cyclic moieties and the alkyl carbon chain do not match any rule potentially responsible for mutagenic/nonmutagenic activity listed in the ToxRead software. Conversely, the bromine atom linked to an aliphatic carbon ring corresponds to two ToxRead/ToxDelta SAs both referring to bromo-/halo-ethyl moieties with different levels of specificity and a prevalence of mutagenic activity of 71% and 67%, respectively. The identified SA extracted by ToxDelta in the source molecule are as following:

Name: MNM16

Description: IRFMN alert n. 16 for Mutagenicity, defined by the SMARTS: [Cl,Br,I][C;H1;D3][$([C;H3;D1]),$([C;H2;D2][C,O,N,S,Cl,Br,I])]

Experimental accuracy: 0.71

Fisher test p-value: 0.00704

Name: SM93

Description: Sarpy alert n. 93 for Mutagenicity, defined by the SMARTS: C(C)Br

Experimental accuracy: 0.67

Fisher test p-value: 0.0036

These rules, which are present in the source molecule but not in the target chemical, give a first indication of different toxicological profiles for these chemicals.

Considering that no other significant mutagenicity SA is identified in the structure of **Mepitiostane**, the user may conclude that although the similarity index between these molecules is high (0.774), the similarity and dissimilarity investigation do not provide evidence to the possibility of assigning the property of the source molecule to the target. Not being available any other SA inside the MCS or the dissimilar fragments extracted from the target after the comparison with the source, the user may conclude that **Mepitiostane** is not mutagenic. Indeed, we already have the Ames test value for **Mepitiostane**, which is negative for this endpoint.

Considering the two case studies, it is notable that sometimes the identified dissimilar substructure is not an entire SA. In many cases the dissimilarity substructures are fractions of the whole SA (i.e. a rule which is present in the database of mutagenicity or any other toxicity endpoint SAs), and the remaining of the SA are in the MCS. This concern has been solved by ToxRead. In fact, the examination of ToxDelta of the dissimilarities happens after the visualization of ToxRead results.

ToxRead outcome comprises all the existing SAs that are matched with the target molecule and are in common between the target molecule and a set of structurally similar molecules.

Table 12. The two case studies: Case study 1) target molecule: Diazepam, source molecule: Flunitrazepam; Case study 2) target molecule: cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-, source molecule: mepitiostane, and the results of ToxDelta: maximum common substructure and dissimilar fragments.

		Molecules	MCS*	Dissimilar fragments
Case study 1	Target			Cl–CH_3 (Aromatic ring)
	Source			F–CH_3 (Aromatic ring); O^-–N^+(=O)–CH_3 (Aromatic ring)
Case study 2	Target			(Aliphatic ring); (Aliphatic ring)
	Source			(Aliphatic ring); (Aliphatic ring); Br–CH_3 (Aliphatic ring)

* Maximum common substructure

4.3 Part 3 – Metal Oxide NMs Genotoxicity Model

The aim of this study is the collection of all the metal oxide NMs with genotoxicity data available in reliable peer reviews and the development of computational grouping model for the created dataset. The prepared dataset is based on evaluation of genotoxicity results from the comet assay results, gathered by the authors[128].

The availability of a relatively high number of publications of Comet assay data for the assessment of genotoxicity of NM, had been the reason of choosing this genotoxicity test in our investigation.

To conclude an accurate result from the various number of studies conducted on 16 metal oxide NMs in the field of genotoxicity by Comet assay, we needed to examine the significant factors that could affect the reliability of the data. Moreover, this scheme made maximum use of pre-existing data from *in vitro* methods, for (Q)SAR modelling.

4.3.1 Data Quality Assessment

Table S1.A in Appendices, contains the results of both physico-chemical properties and Comet assay adherence assessment. Each datum point corresponds to the summary of results from the Comet assay for a nano metal oxide (or silica), with a common core chemical composition and a unique size/ size range.

As part of this assessment, we extracted data for the size of various metal oxide NMs and provided in Table S1.B (Appendices). In a few publications, size was reported as nominal size as provided by the suppliers whilst other authors measured the size with one or more tools including (TEM, SEM, etc.).

Information on the crystallography is also included in this table. Metal oxides that had different crystallographic properties in the assessed data were TiO2, and SiO2. The anatase or rutile forms of TiO2 and amorphous or crystalline forms of SiO2 were tested.

Physico-chemcial properties of the NMs such as size, shape, charge, and surface coating, and various components present in the medium, such as serum proteins that dispersion and stability of NMs depend on them, influence the assay results. Considering the importance of these factors in Comet assay results, it is essential to characterize NMs in the relevant medium and to use the appropriate treatment conditions.

Recent studies in the field of NMs show that significant size-dependent changes in NMs' properties happens in the NM smaller than 5 nm [150–152]. NMs with sizes between 15 and 90 nm do not show a meaningful correlation between the factor of size and activity. On the other hand, all nanopowders when suspended in water resulted in same sized aggregated particles, regardless of their initial size [153]. Thus, we decided to exclude NMs smaller than 15 and bigger than 95 nm from the consecutive data analysis.

In data quality assessment section the aim was to examine the already existing data and establish a data set of metal oxide NMs with their assigned genotoxicity properties, based on the minimum criteria. This dataset was aimed to be used for (Q)SAR modelling purpose [154]. It is important to notice that until the current time there are no official guidelines for the *in vitro* Comet assay [155]. After assessing the quality of the tests presented in the literature, we propose a scheme employing a WOE approach [133]. This approach is an appropriate solution to make a conclusion for the hundreds of reports that have been published without a concrete outcome [156]. Finally, a binary classification of genotoxic or nongenotoxic has been assigned to each metal core composition.

We established some criteria to make a possible comparison of data obtained from the reports of different laboratories and studies. Table 5 (page 69) shows the existing trend for genotoxicity of each type of the metal oxide NM as identified by the Comet assay. The differences of Comet assays methods in different laboratories have been investigated in past [157–163]. In our approach, we selected the results with sufficient number of necessary factors for the performed test. These important factors of the Comet assays make their results more reliable. We have undertaken two phases to evaluate the experimental data: those criteria to be respected for a reliable *in vitro* Comet assay and a WOE approach applied to the reports that have met the criteria established. Previously, Huk et al.[154] establish the minimum considerations in performing the *in vitro* Comet assay experimentally for NMs. These considerations are listed in Table S1.A (Appendices). For instance, data points that are not associated with the three important physico-chemical characterisations (i.e. chemical composition and purity, surface area and particle size/size distribution) are known to be less reliable comparing to the NMs for which this kind of

information was provided. The data lines without any physico-chemical properties of tested metal oxide NMs, were excluded from the list of data. Therefore, the data points assessed for *in vitro* Comet assay minimum protocol were satisfactory from this point of view.

After applying the established criteria to the data points, 48 data belonged to the "high reliability" class of which 18 data points had also uptake analysis. 8 data points belonged to the "moderate reliability" class, of which 2 had an uptake analysis. 4 data points belonged to the "low reliability" class, of which 1 had an uptake analysis. Table 13 (page 91) shows the assignment of the reliability of Comet assays based on the criteria list defined by Huk et al. [154]. Details of this analysis are reported in Table S1.A (Appendices). Table S4 (Appendices) reports the number of total studies evaluated for each metal oxide core (all sizes) together with the number of studies with size range 5-100 nm, the number of studies in each reliability class and the overall genotoxicity for each metal oxide nanoparticles.

4.3.2 Quantum Mechanical Descriptor Calculations

The size of the NPs under investigation in this study is 15-90 nm in the laboratories' analysis. Calculation of the quantum-mechanical descriptors of these NMs is not possible, since the systems are too large. For this reason we decided to simplify the structural models used for the calculation of the descriptors. Smaller metal oxide clusters of the same size have been considered for the calculation of the descriptors and one of the descriptors was calculated on the basis of the characteristic of the considered NMs. In the current study, we adopted the same method used by Puzyn et al.[164] for a predictive cytotoxicity model. In our study we used the same 10 metal oxides that Puzyn et al. used in their cytotoxicity dataset. Puzyn et al. utilized 16 NMs in genotoxicity and 17 NMs in cytotoxicity dataset. The same concepts in considering experimental results performed on different sizes from 15 to 90 nm has been adopted in the present study. In addition, Puzyn et al. have given a strong justification for the use of the size rage of 15-90 nm outcomes in a singular modelling approach. In precedent studies it is shown that genotoxicity of many NMs is directly related to oxidative stress (by elevated ROS levels, reduced antioxidant levels and increased lipid peroxidation) and subsequent inflammation

(leading to apoptosis)[165,166]. Hence, the descriptors that were selected by the machine learning approaches were evaluated to illustrate their role in the oxidative stress process.

4.3.3 Recursive Partitioning and Regression Trees

In order to maintain the simplicity of the model, we built a simple classification tree on the data set of 16 metal oxide NPs. Table S3 (Appendices) reports the scaled values on the dataset used for the classification model. In the developed model tree descriptors have been used and the NMs are classified into two classes: genotoxic and nongenotoxic. Figure 8 (page 89) shows the tree representation of the developed model. Since the model is fitted with scaled data, for making predictions the new observations should be scaled according to the scale attributes (Table S3, Appendices) of the dataset used to build the model. The thresholds used in this model are as following:

First split:

- NPs with HF < -1.145 (original value: -5199.7) are nongenotoxic; the rest of the NPs will be processed in the second split.

Second split:

- NPs with MW < 0.3601 (original value: 1492.2) are nongenotoxic; the rest of the NPs will be processed in the third split.

Third split:

- NP with COSMO-SA >= 1.176 (original value: 960.9) are nongenotoxic; and the rest are genotoxic

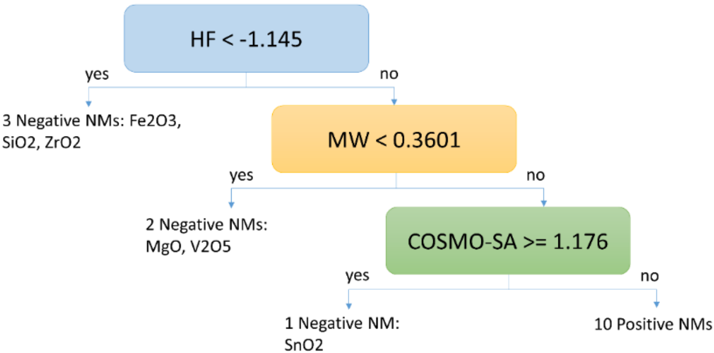


Figure 8. The 'rpart' classification tree model on the data set of 16 metal oxide nanomaterials.

All the eleven descriptors reported in Table 4 (page 68) are given as input to the rpart algorithm. The method is set to classification, as the endpoint is a binary value (i.e. genotoxic and nongenotoxic). Three control parameters are set in the command line of building the model. "Minsplit" that is the minimum number of observations in a node for which the routine will even try to compute a split is set to 3, "minbucket", which is the minimum number of observations in a terminal node is set to 1, and "cp" that is the threshold complexity parameter is set to 0.001. The resultant model successfully separated the sixteen NPs into two groups using three quantum-chemcial descriptors: heat of formation (HF), molecular weight (MW), and surface area of the oxide cluster based on conductor-like screening model (COSMO-SA). The other eight descriptors did not appear in the final model.

The selected descriptors very well conform to overall toxicity model for metal oxide NM, while this selection also can be beneficial for explanation of genotoxicity mechanisms of action. The results show that NM with small MW and HF are less genotoxic, comparing to heavier metal oxide NM. Enthalpy of formation or heat of formation is the only structural descriptor used in the nano-(Q)SAR model developed by Puzyn et al.[164]. From an initial set of 12 quantum-chemically calculated structural descriptors, ΔH_{Me+} representing the enthalpy of formation of a gaseous cation with the same oxidation state as that in the MO structure was used to establish the linear equation of the model. HF is associated with the stability of the NM. Possibly, NMs, which are less stable, release ions more easily, and this increase the effect. COSMO solvent accessible surface area (COSMO-SA) is a continuous surface of the molecule, which can be reached by the centre of charge of a solvent molecule. The correlation between the lower COSMO-SA values and the higher DNA damage can be explained by the fact that the metallic ion has to be supplied via solution to DNA. The solubility and the dissolution rate of the metal oxides are essential for the metal supply. For the same type of metal ion (Fe), more metal ions are accessible in Fe3O4 comparing to Fe2O3. Assuming the surface area is equal, different number of oxygen, the metal type and the crystal structure (number of oxygens for each atom of metal), play an important role. This information are encoded in COSMO-SA. In other words, two different metal oxides having equal surface area can have different number of metal ions on

the surface depending on the length of the inter-intra molecular bonds, volume/molecular weight of the metal ion, and the proportion of the number of oxygens to the number of metals. Because of the small size of the dataset in our study in addition to the classification endpoint, we decided to maximize the amount of data in the process of model training and consequently there were no data considered as test set. Nevertheless, all the principles of the OECD for the validation of (Q)SAR models [73] have been fulfilled in our new model, except the model validation for stability and predictivity. In fact, the developed model has a defined endpoint: genotoxicity, an unambiguous algorithm, and an applicability domain. The most important goal of developing a new (Q)SAR model for the provided data set is to evaluate the correlation of each quantum-mechanical descriptor with the genotoxicity endpoint of the NMs. Indeed, feature selection is a crucial phase of (Q)SAR modelling. The developed model reveals the most significant descriptors related to the genotoxicity risk assessment of the NMs from a quantum-mechanical point of view.

Table 13. Assignment of the reliability of *in vitro* Comet assays based on the criteria defined in Huk et al. The assignment questions were treated in a "yes" or "no" fashion. In a weight of evidence approach, data points that presented 1st class property were used to assign the genotoxic or non-genotoxic fate to the metal oxide nanomaterials with the same chemical core composition.

Comet assay main principals to follow for obtaining reliable test results	1st class data	2nd class data	3rd class data	Unreliable data
Exposure to the nano metal oxides expressed in at least one of the units mentioned in Table 6 (page 70)	yes	yes	yes	no
Cytotoxicity tests performed? (results were in a range from nontoxic to around 80% viability)	yes	no	no	no
At least exposure for (3h) treatment time	yes	yes	no	no
Performed trend test for dose-response relationship?	yes	no	no	no
Performed comparison between treated samples and controls or both?	yes	yes	yes	no
Information on uptake (demonstrated cellular uptake?)	yes (prioritized results) and no	yes (prioritized results) and no	yes (prioritized results) and no	no

4.4 Weight of Evidence Approach in In Silico Models for the Mutagenicity Assessment of Chemicals

There can be two methodologies for the assessment and integration of the prediction results of the *in silico* models. In the WOE section of the present dissertation, to develop a conclusion in a WOE approach: i) first the (Q)SAR and read-across results as reported by the prediction models are documented and integrated, considering the reliability, relevance and consistency of each prediction, ii) then the structurally similar compounds indicated by the *in silico* software, as most similar chemicals to the target chemical under investigation are explored in order to evaluate the relevance and the reliability of the similar chemical to be used in a read-across way. We employed the both methodologies to the pieces of evidence. First we considered the prediction results of each individual *in silico* model with their corresponding applicability domain index as a measure of reliability. Then the similar chemicals indicated by different *in silico* models are taken into consideration for a further investigation to assess their relevance in terms of read-across. This assessment is based on their structure characterization or the presence or absence of active or inactive mutagenic rules and the similarities and dissimilarities between the target and the source molecules. At the end, the consistency of the various results has been evaluated together with their reliability and relevance to develop a conclusion for the mutagenicity effect of each drug. The results of the two case studies are presented in Chapter 6.

CHAPTER 5

5. Conclusions

5.1 Part 1- Carcinogenicity Models

Based on SARpy as an automatic SA extraction tool, we developed two carcinogenicity models from two different training sets. The ANTARES learning set contained rodent bioassay carcinogenicity data of 1543 compounds, while ISSCAN-CGS dataset consisted of 986 chemical substances with human-based assessments and data retrieved from different types of assays. We thoroughly evaluated the predicitivity of each model, this was evaluated on its related test set and additionally on an external test set composed of 258 chemical compounds obtained from the ECHA inventory. These two newly developed models are implemented in the VEGA platform and are freely accessible to all users.

The already existing carcinogenicity models in the VEGA platform [167] before adding the two developed models (ANTARES and ISSCAN-CGX) were CAESAR and ISS. The CAESAR model is built as a CP ANN model. The neural network output consists of two values labeled as Positive and Non-Positive. The CP ANN uses twelve molecular descriptors. The ISS model is indeed the Toxtree carcinogenicity module version 2.6 [52] introduced by Benigni et al. [43]. When at least one carcinogen rule is matching with the target compound, "carcinogen" prediction is given; otherwise, the prediction will be "non-carcinogen". The carcinogenicity developed models are the only structure-based models of the VEGA platform, except the ISS model which is indeed the implementation of Toxtree carcinogenicity module in this platform. The SAs of the ISS model are exactly the same as SAs of the Benigni and Bossa ruleset [43].

Automated discovery of SAs associated with toxicology has been made great progresses due to the evolution of data mining tools. The statistically-based methods for the identification of new SAs are helpful in improving the already existing rule sets. While the most known carcinogenicity rule sets [43] are collected on the basis of human expert judgement, the SAs identified in our study are extracted by SARpy with no *a priori* knowledge about the MoA of the chemicals. This approach highlighted some new clues about genotoxic and non-genotoxic

SAs. Some primary analyses have been provided on the SA lists; chemical classes of the identified SAs have been evaluated, however, further study for the new SAs should be performed.

In addition to the newly discovered SAs in the present study, the results obtained by the SARpy SAs extraction based on the analysis on the ANTARES and ISSCAN-CGX data sets are completely in line with the SAs presented by Kazius et al.[46].

Furthermore, the models are developed on the basis of two learning sets with different carcinogenicity data from the point of view of origin and provenance. Concerning the training sets with variant carcinogenicity data assessed within different properties, each set of the extracted SAs constituted a purpose oriented model. The user may consider the results of the model with more realistic predictions (ISSCAN-CGX) or the one with more conservative assessments (ANTARES).

From a decision-making point of view, the most logical approach is combining the evidences obtained from different sources of information such as (Q)SAR model predictions, *in vitro* and *in vivo* test results. Scientists instead of accepting a singular judgement or prediction from one source of information, may acquire a WOE approach and consider more methodologies before determining the level of toxicity of a target substance. An example of the latter approach is implemented for mutagenicity (Ames test) endpoint, in the VEGA platform, in which the results of different models are combined and the output is based on different existing models [167]. These two developed models for carcinogenicity are also implemented in VEGA where other models for the same endpoint are available. This implementation increases also the possibility of performing a similar activity to make a conclusion.

Finally, the SAs explored in this study will be used for the construction of the carcinogenicity ruleset in ToxRead (http://www.toxgate.eu), a platform that uses set of rules for different endpoints to filter and select similar compounds and assist the user in performing read-across studies [122,123].

5.2 Part 2- ToxDelta

ToxDelta is a complementary tool to be used in a parallel way together with the other prediction tools. It can be used alone and also is aimed to match some important features of ToxRead. ToxDelta addresses the dissimilarities between two similar molecules, and it does not make a conclusion on the evidences found in the dissimilar fragments of two substances. The conclusion on the outcome of the prediction may be accomplished by other prediction or read-across tools such as ToxRead. The main advantage of this new read-across tool is the emphasis on the differences in addition to the similarities and the resembling properties between two structurally similar molecules. In other words, it exploits the adverse effects of the dissimilar fragments that may trigger the toxicological properties or biological activities of the chemicals. ToxDelta analyses the modulations of the effects related to the presence of the specific fragments in one of the two structures under investigation. ToxDelta executes in a "local" way. This functionality is important to evaluate the metabolites and the impurities related to a target compound in a comparative approach to the parent compound. Two important fields of the application of this strategy are impurities in pharmaceutics and pesticides. The FDA has provided a guideline for industry about the mutagenicity of the pharmaceutical impurities [168] that describes a practical framework for identification and control of the identified mutagenic impurities in order to limit potential carcinogenic risk. Another appropriate field of application for this tool is in pesticides assessment. The EFSA has discussed the use of *in silico* models for the evaluation of the effects of metabolites of pesticides. ToxDelta's methodology is useful in the toxicity assessment of pesticides, biocides and pharmaceuticals. The experimental toxicity properties of the parent compound is requested by the regulatory bodies and ToxDelta is able to provide this information. In these cases data for the parent compound is available and user requires the possible increase of effect caused by an impurity in the structure of a chemical instead of the absolute effect of the related compound. If the toxicity level of the impurities is similar to the parent compound, there must not be differences in the regulations and laws related to the target compound. Conversely, if the impurities have potential hazard effect the compound under the investigation must go through more analysis for the hazard assessment. To resolve these problems, local tools that are able to measure the relative increase or decrease of the

effects are more accurate than absolute *de novo* predictions. Despite the widespread use of read-across tools, still the acceptance of the dossiers based on read-across approaches is not straightforward. A detailed documentation is required to be provided by the expert. The main concept of read-across is constructed by the analysis of two (or more) substances, the target compound, with missing data, and the source compound(s) which is assumed to represent the properties of the target compound. The already existing read-across tools to this moment are all focused on the similarities between the target and the source compounds. The main idea is the higher the similarity is, the greater is the likelihood that the two compounds share the same biological properties/activities. Indeed, the authorities often discuss the fact that even minor differences may provoke crucial change in the properties of the substances. In order to complement the existing read-across tools, we put emphasis on the differences between two structurally similar compounds, introducing ToxDelta.

Another noticeable difference between ToxRead and ToxDelta with the other read-across programs, is that they do not contain exclusively active SAs, but also inactive SAs. This allows the examination of any positive or negative modulations of the effect. To each SA statistical characterizations are assigned. These statistical values show the accuracy and p-value of the SAs and are calculated based on the number of chemicals containing that SA, and the prevalence of the toxic or non-toxic category. Consequently, the tool provides not only the SAs present in the compounds, but also the statistical significance of the association of the found SA to a certain effect. ToxRead contains data related to mutagenicity and BCF endpoints and permits the user to move in different levels of reasoning in a read-across approach. ToxDelta offers additional focus on all the dissimilar fragment that may affect the properties and the activities of the molecules.

Currently, a beta version of ToxDelta is freely available on the VEGA platform (https://www.vegahub.eu/) and the toxicity endpoint for which this tool can be used is mutagenicity. Other endpoints will be added to the software in the future.

5.3 Part 3- Metal Oxide NMs Genotoxicity Model

NMs are complicated and exploring the relationships between their properties and their toxicity is challenging comparing to the small organic substances. This study is the first attempt in the field of computational nanotoxicology for modelling the genotoxicity of metal oxide NMs. Even though sufficient experimental data on genotoxicity of metal oxide NMs are available, there is no reliable model (or theory) for the prediction of genotoxicity of these NMs. In the present study, (Q)SAR modelling has been conducted based on experimental data obtained from a wide range of tests done on metal oxide NMs by *in vivo* and *in vitro* methods, mainly the *in vitro* Comet assay. We applied data quality assessment techniques on the experimental data based on *in vitro* models to create a reliable data base for use in computational modelling.

The results of our WOE assessments confirm that it is important to fill the gaps of physico-chemical characteristics of NMs used in *in vitro* Comet assay. In our analysis the range of size that has been mostly covered in experimental testing has been reported. We aimed at ranking the potential genotoxic category by material class, focusing on the within and between class variability. Within our investigation, we examined the role of the size of the same metal composition chemical core (size range or 1 to 90 nm) in genotoxicity effect of these chemicals as measured by Comet assay, and the possible statistical inferences that can be obtained from those data.

While there are numerous progresses in the field of traditional (Q)SAR analysis, nano-(Q)SAR modelling is still at its primitive phase, due to the lack of sufficient knowledge about the measurements and modelling standards. The most problematic part of nano-(Q)SAR modelling is defining a series of consensus characterizations for the toxicity tests. The standardization of NMs characteristics and test methodologies is a great step towards the realization of successful nano-(Q)SAR models.

In this study, for the first time we introduce a genotoxicity model which relates the experimental genotoxicity property of a set of metal oxide NMs to the quantum-mechanical calculated descriptors of these NMs. We successfully built a classification nano-(Q)SAR model based on a simple tree modelling approach. The aim of this model is to identify the most significant quantum-mechanical descriptors of the NMs that affect the genotoxicity properties assigned to

each NM based on a WOE study applied to a number of peer review resources. To the best of our knowledge, this study is the first example of NMs genotoxicity modelling by nano-(Q)SAR approach. In consequence of the small number of samples in the data set and the classification endpoint, we decided to use all the available data as training set in order to study the variable importance, rather than focusing on prediction ability of the model. Even though the model is simple and does the classification in a straightforward mode by using only three quantum-chemical descriptors, it is still based on a small data set and the model validation still needs to be accomplished using a test set. Although the initial findings are encouraging, there is a strong need to verify and validate the results in order to make them acknowledged by regulatory bodies and users. Concerning the restraints, the developed model demonstrates a high potential of current chemo-informatic approaches for toxicological assessment of various metal oxide NMs. The design and manufacture of safer NMs require detailed analysis. The use of this model during the early stages of risk assessment can be very helpful to prioritize the NMs, which may impose adverse effects on human health.

Although the mechanisms of nanoparticles genotoxicity are still not fully discovered but direct DNA damage and oxidative stress are considered important [169,159]. Direct DNA damage mechanism is assumed to be more nano-specific because small nanoparticles may reach the nucleus through the nuclear pore complexes [170]. However, the observation of larger nanoparticles in the nucleus hints that larger nanoparticles may get access to the DNA in dividing cells during the nuclear membrane dissemblance [171]. Oxidative stress is induced by overproduction of ROS, resulting in the loss of normal physiological redox-regulated functions in the cells. This triggers DNA damage, unregulated cell signalling, change in cell motility, cytotoxicity, apoptosis, and cancer initiation [172]. The relationship between DNA adducts and oxidation-induced DNA fragmentation and exposure to metal oxide nanoparticles is confirmed in numerous studies [173-177]. Huang *et al.* [178] reported a detailed description of the genotoxic mechanisms of action of metal oxide nanoparticles.

Enhancement of systematic risk assessment for NM is one of the main emphasis of the topic-related (EU)-funded projects. The ongoing “Nanosafety Cluster” aims at identifying key areas for further research on risk assessment procedures for NM. The NanoSafety Cluster Working

Group 10 outcome was concern-driven integrated approaches for the (eco)-toxicological testing and assessment of NM [179]. A set of tiers using standardized protocols for preparation and testing was introduced. Tier 1 included determining physico-chemical properties, non-testing methods such as QSAR and evaluating existing data. Tier 2 consisted of performing a limited set of *in vitro* and in vivo tests that are used to clarify the known risks of a relative concern or to highlight the need for performing further tests. A concern-driven guidance for investigating potential risks of NM is based on the idea of focusing research on NM that may induce some concerns based on exposure, use and already available toxicological information driven form non-testing methods ((Q)SAR, pharmacokinetic modelling and read-across). A testing strategy should consider the possibility to apply "read-across" methodology, to omit the "non-necessary" tests based on the relative category of a NM. The aim is to improve the risk assessment strategy in order to require less testing whenever the available information is sufficient to reach a conclusion in a decision-making process.

Based on the factors determined that contribute to the genotoxicity of metal oxide NMs further studies will be performed to determine structural features which may help to derive mode of action knowledge from the data, i.e. prove a key mechanism that can describe a DNA damage.

5.4 Overall Conclusion

The growing number of chemicals requring risk assessment, envisions increased efficiency in toxicity testing and the way toxicity testing is currently conducted may need major changes. In fact, acquiring a full set of toxicity results needed for regulatory bodies and decision-making procedures at the same pace these chemicals are introduced to the market, is becoming ever more challenging. Several approaches have been introduced in recent years, including the use of robotic high-throughput screening and computational toxicology studies to overcome this issue while decreasing the animal usage and increasing the required time for testing. Additionally, relibale non-testing methods, including (Q)SAR and read-across are becoming more and more acknowledged by risk assessors and regulators. The OECD experience with non-testing methods illustrates that grouping chemicals into categories and filling data gaps by read-across, interplolation or extrapolation are a winning strategy. Grouping approach relies on the assumption that not all the chemicals belonging to a group need to be tested for all toxic effect, and data gaps can be filled by means of read-across and grouping, saving “unnecessary” *in vivo* or *in vitro* tests. The flexibility and the transparency of the read-across approach make it a successful tool in many toxicological fields, especially regulatory decision-making. Consequently, these tools contribute to the realization of the Russell and Burch’s “3Rs principles” (Replacement, Reduction and Refinement) of animal use in toxicological studies. The use of the proposed WOE framework is illustrated by two drugs as case studies for mutagenicity assessment (Chapter 6). ToxDelta and ToxRead as two read-across tools are powerful means for the exploration of the active and inactive SAs present in the target molecule and the similar chemicals used as sources in read-across terms. These new tools help the assessors to overcome the shortcomings related to the interpretation of the results of *in silico* prediction models. The results show that the combination of the methodologies: (Q)SAR and read-across methods, with ToxRead and ToxDelta develops more interpretable data that can be utilized as lines of evidence in a WOE approach. The structure-based carcinogenicity models introduced in the first part and the new read-across tool, ToxDelta introduce in the second part, are completely in line with the aim of improving the current methodologies in WOE approaches.

In the third part of the study, the new genotoxicity model developed for the metal oxide NPs provides assistance for identifying and prioritizing the NMs which pose toxic effects to human health and environment. Safety assessment of NM and their modifications (shape, size, surface, coating, etc.) needs a full-blown testing program for each NM. This leads to a huge amount of testing with their relative costs and testing time. Some hazard information of NM can be deduced from the similar bulk materials or similar NM. In general, read-across and category approaches are used to predict properties and/or biological effects of chemicals. This is a way to fill the data gaps to characterize the adverse effects of NM. The WOE approach is also an effective tool for the integration of the conflicting and different results collected from the literature. From a regulatory point of view, the European projects which aim at introducing new strategies for further investigation on NM toxicity, point at a concern-driven guidance. This concern-driven guidance for investigating potential risks of NM is based on the idea of focusing research on materials that may induce some concerns based on exposure, use and toxicological information driven form non-testing methods ((Q)SAR, pharmacokinetic modelling and read-across). Whenever possible, a testing strategy should consider the application of "read-across" methodology, to reduce the number of assays based on the potential risk associated with a NM. The aim is to improve the risk assessment strategy of NM in decision-making processes, and we believe that our classification tool can be used as an effective QSAR model to prioritize the metal oxide NM with high concern.

The main goal of (Q)SAR and read across studies is to improve the risk assessment strategy in order to require less testing whenever developing new data is feasible. Using data integration approaches can help scientists and regulators in decision-making processes, and enable them to reach conclusions. The results obtained during my studies and presented in this dissertation are useful progresses in the field of structure-based genotoxicity non-testing methods. In particular, new approaches to read-across studies, achieved and implemented during my Ph.D studies, provide new scientific contributions to a transparent and structured framework for chemicals and NM risk assessment.

CHAPTER 6

6. Example of the Use of Non-Testing Methods within a Weight of Evidence Framework

In this section two case studies are selected for demonstrating the use of non-testing methods within a WOE approach. The prediction results for mutagenicity effect of a target compound are considered as pieces of evidence, while the objective is integrating these results to reach a conclusion on mutagenicity of the chemical under investigation. These pieces of evidence form a line of evidence to be used in further investigations together with other types of lines of evidence to help the assessors to support an adequate answer to a toxicity question. Two drugs are selected for the present practice: Valproic acid and Diclofenac.

The evaluation and integration of the *in silico* model results in order to develop a conclusion in a WOE approach can be performed in two ways: i) first evaluating the (Q)SAR and read-across results as reported by the prediction models and integrating the results, considering the reliability, relevance and consistency of each prediction, ii) analysing the structurally similar compounds indicated by the *in silico* software, as most similar chemicals to the target chemical under investigation in order to evaluate the relevance and the reliability of the similar chemical to be used in a read-across way. In the present study, we employed both methodologies to the pieces of evidence. First we considered the prediction results of each individual *in silico* model with their corresponding applicability domain index as a measure of reliability. Then the similar chemicals indicated by different *in silico* models are taken into consideration for a further investigation to assess their relevance in terms of read-across. This assessment is based on their structure characterization or the presence or absence of active or inactive mutagenic rules and the similarities and dissimilarities between the target and the source molecules. At the end the consistency of the various results has been evaluated, together with their reliability and relevance to develop a conclusion for the mutagenicity effect of each drug.

6.1 First Case Study: Valproic Acid

For the first case study we chose Valproic acid (Figure 9-page 103) as a target molecule for its mutagenicity assessment (as assessed through bacterial reverse mutation test) by non-testing methods, such as (Q)SAR models and read-across. For the estimation of mutagenicity we opted two platforms: VEGA (http://www.vega-qsar.eu) and T.E.S.T. (http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST). The VEGA platform contains four mutagenicity models (CAESAR, SARpy, ISS and KNN) and a consensus model which makes a conclusion on the basis of the four models predictions. The T.E.S.T. platform encloses three mutagenicity methods (Hierarchical, FDA and Nearest Neighbour). The Hierarchical and FDA methods are (Q)SAR structure-based methods, while Nearest neighbour is a read-across model. The mutagenicity property of the molecule under investigation is also assessed by ToxRead (http://www.toxread.eu/), to provide additional support for the predictions. Further T.E.S.T.'s outcome provides a few examples of similar molecules to the target molecule along with their similarity index, experimental and predicted values to be examined by the user. In the present practice, we applied both methods to our investigation in a WOE approach. Here we reported some of these similar examples presented by T.E.S.T. and evaluated their structural similarity and dissimilarities and the effect of the dissimilar fragments in the properties of the source molecules comparing with the target.

Figure 9. Valproic acid chemical structure (CAS number: 99-66-1)

Table 14. Summary of the prediction results of non-testing models for Valproic acid

Software	Model/method	Experimental value	Results	Predicted value	Applicability Domain Index
VEGA	CONSENSUS		NON-Mutagenic		Non-mutagenic score: 0.9
	CAESAR	N/A**	NON-Mutagenic		0.98
	SARpy/IRFMN	N/A	NON-Mutagenic		0.98
	ISS	N/A	NON-Mutagenic		0.90
	KNN/Read-across	N/A	NON-Mutagenic		0.96
T.E.S.T.*	Consensus		NON-Mutagenic	-0.04	Internally checked
	Hierarchical	Non-mutagenic	NON-Mutagenic	-0.01	Internally checked
	FDA	Non-mutagenic	NON-Mutagenic	-0.1	Internally checked
	Nearest neighbour	Non-mutagenic	NON-Mutagenic	0	Internally checked
ToxRead	Read-across		NON-Mutagenic		N/A

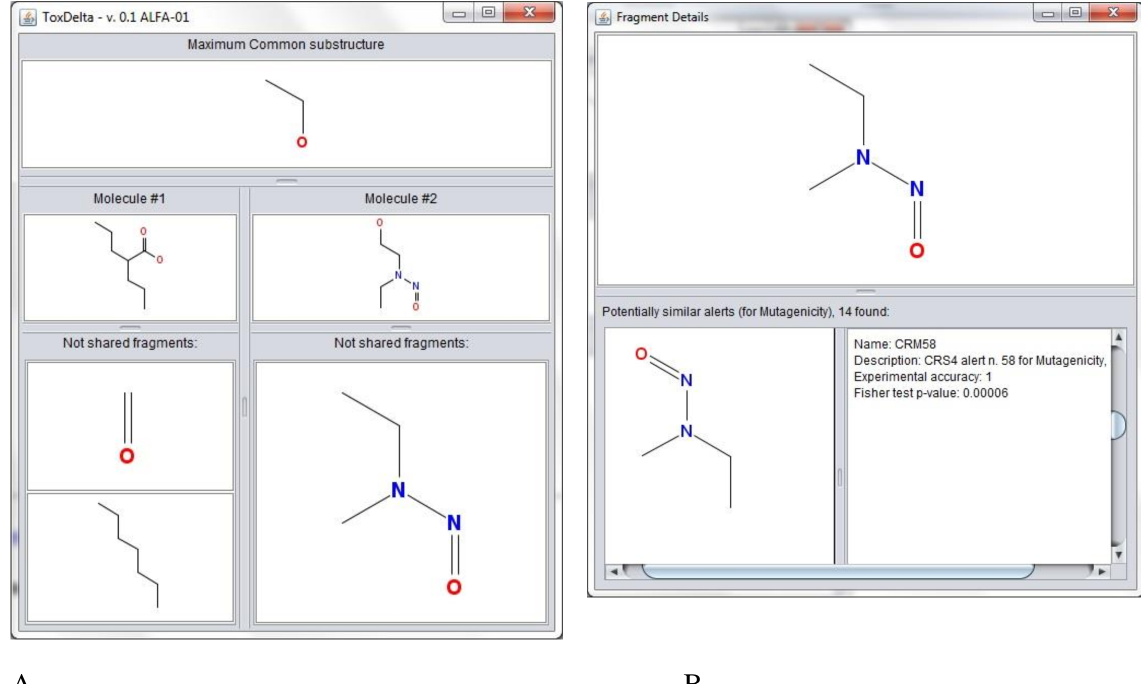
*the test chemical was present in the training set.
**Not available

Table 14 (page 104) shows the predictions of all models. All the models of VEGA and all the methods of T.E.S.T. predicted the substance as non-mutagenic. The applicability domain indices for each model in present in the table and they illustrate the level of reliability for each prediction. All the models agree on the non-mutagenic effect of the chemical. These results are based on experimental values, structurally similar compounds and the results of (Q)SAR models. The target substance was present in the training set of T.E.S.T. as non-mutagenic. Among with the predictions of all the models and the consensus method of T.E.S.T. also a

number of similar substances in the training set and external test set are reported in the outcome panel. For sake of example we reported four similar substances identified by T.E.S.T. in Table 15 (page 106). For each similar substance the corresponding similarity coefficient, experimental value and predicted values are reported. Oryzalin metabolite and 2-[Ethyl(nitroso)amino]ethanol are two similar substances with high similarity coefficient (0.74 and 0.81, respectively) which reports positive mutagenicity effect for the experimental and also predicted values. The reason for the conflicting mutagenicity result for these two molecules is evaluated by ToxDelta. 2-[Ethyl(nitroso)amino]ethanol (CAS 13147-25-6) (mutagenic) is selected to be compared to Valproic acid (non-mutagenic). Figure 10 (page 107) reports the outcome of ToxDelta for the MCS extraction and dissimilarities identification. After the subtraction of the MCS form the target and the source substances, the dissimilar substructures are shown in the outcome panel. The two dissimilar substructures extracted from the target substance do not present any mutagenic potentiality, but the dissimilar substructure extracted from the source molecule is known to be an active mutagenicity rule in the CRS4 mutagenicity rule base with accuracy=1.

Table 15. Experimental and prediction values for some examples of similar chemicals to Valproic acid (CAS number: 99-66-1) in the training set and test set of T.E.S.T.

CAS	Structure	Similarity Coefficient	Experimental value	Predicted value
621-64-7		0.81	1.00	0.97
110-58-7		0.79	0.00	-0.02
106-27-4		0.74	0.00	-0.00
13147-25-6		0.74	1.00	0.90



A B

Figure 10. A) ToxDelta outcome for the comparison between Valproic acid (molecule #1) and 2-[Ethyl(nitroso)amino]ethanol (CAS 13147-25-6) (molecule #2) The maximum common substructure is at the top of the panel. The dissimilar substructures are listed below their corresponding molecule. B) The identified dissimilar substructure extracted from the molecule #2 is a mutagenicity structural alert in the CRS4 dataset of mutagenicity ruleset with accuracy=1.

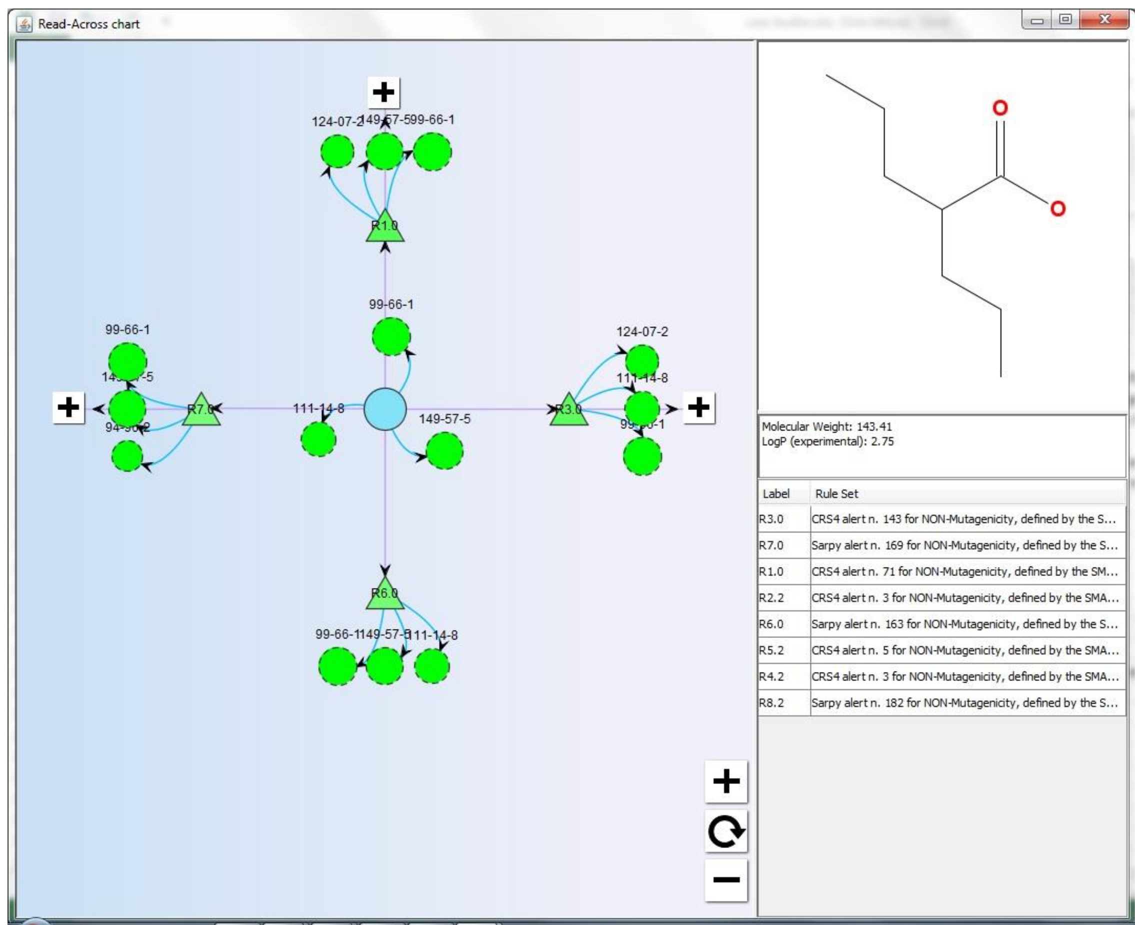


Figure 11. ToxRead chart for the target molecule Valproic acid (CAS number: 99-66-1). The numbers refer to CAS identifiers. Straight arrows link the target chemical to rules, while curved arrows link to chemicals

ToxRead as a read-across tool can integrate the results from (Q)SAR and read-across models in a WOE approach. While some of the similar substances identified by T.E.S.T. are not structurally so similar to Valproic acid, the similar substances identified by ToxRead in the process of read-across are more compatible with the target (Figure 11-page 108). Indeed, the target chemical does not contain any structural rule for mutagenicity. The presence of nitrosamine substructure in the source substances triggers a remarkable difference in biological effect of the molecules form the mutagenicity point of view. This crucial differences and their role in reducing or amplifying toxicity effects of the molecule (not only mutagenicity but also other endpoints) can be precisely established by ToxDelta and can have an important role in the process of decision-making and the expert judgement in terms of WOE. Indeed, risk assessors can make use of this tool in order to minimize or even eliminate the eventual uncertainties

during the assessments with non-testing methods. Considering the differences between Valproic acid and the similar compounds identified by T.E.S.T. we can draw the overall conclusion that the substance under investigation is non-mutagenic. Table 16 (page 109) is the tabular format for summarizing WOE.assessment of Valproic acid in a qualitative way.

Table 16. Optional tabular format for summarizing weight of evidence assessment of Valproic acid

Question		Hazard identification
Assemble the evidence	Select evidence	Nine (Q)SAR models from two *in silico* platforms, a read-across tool, ToxRead and a tool to investigate the dissimilarities between the similar compounds and the target, ToxDelta, are chosen for testing mutagenicity effect of Valproic acid. The used platforms provide both predicted values from several models, and indicate similar substances with experimental values, which can be used for read-across.
	Lines of evidence	All the models estimated the chemical as non-mutagen. The exception were a few examples of similar compounds in the T.E.S.T prediction outcome. To evaluate the relevance of these similar molecules in a read-across approach we used ToxDelta. The dissimilar substructures between the target and the source molecules are investigated. Valproic acid do not contain any mutagenic structural alert, so the similar compounds assigned by T.E.S.T. to the target are not relevant for the assignment of the same property to the target molecule.
Weight the evidence	Methods	VEGA provides applicability domain index that is a sort of quantitative measurement of reliability and values higher than 0.8 are considered more reliable. T.E.S.T. applies a filter to eliminate not reliable predictions. ToxRead indicates the structural alerts found in the target which are associated with the effect and an arbitrary number of similar compounds which share that rule with the target. ToxRead also reveals the structure alerts present in the source compounds, and by using ToxDelta user can check if these active moieties belong to the dissimilarities between the two molecules. For Valproic acid this further evaluation is performed on the similar compounds identified by T.E.S.T. to check the relevance and the reliability of the similar compounds to be used in terms of read-across.

	Results	All the predictions obtained from VEGA and T.E.S.T.are reliable in terms of applicability domain index. No structural alert in the structure of Valproic acid is identified by ToxRead. A further evaluation is performed on the similar compounds suggested by T.E.S.T. with contrary mutagenic effect using ToxDelta to check the relevance of the similar compounds. The results showed that the potential structure alert found in the source compound was not present in Valproic acid, and therefore is not relevant.
Integrate the evidence	Methods	Considering the reliability and relevance of each estimation obtained from individual mutagenicity prediction models, the results are integrated, together with the consistency of the predictions. These results can be used by an expert judgement to reach a conclusion on the probability of mutagenicity effect.
	Results	All the *in silico* methods used in this practice are in concordance with high level of reliability and relevance. Considering the evidence obtained from these predictions, it can be concluded by expert judgement that the target compound is not mutagenic.

6.2 Second Case Study: Diclofenac

Figure 12. Diclofenac, CAS number: 15307-86-5

The second case study presented in this dissertation is Diclofenac (Figure 12-page 111) (CAS number: 15307-86-5). Herein the use of WOE approach when information is derived from non-testing methods, such as (Q)SAR and read-across is described. For the estimation of mutagenicity we used two *in silico* platforms: VEGA (http://www.vega-qsar.eu) and T.E.S.T. (http://www.epa.gov/nrmrl/std/qsar/qsar.html#TEST). Also ToxRead (http://www.toxread.eu) as a read-across tool is used for the further investigation of the target molecule. Table 17 (page 112) shows all the predictions of the (Q)SAR models used in this study. All the VEGA models estimate the compound as non-mutagenic with applicability domain index higher than 0.77, except KNN/read-across with applicability domain equal to 0.65. The VEGA Consensus model predicts the target substance as non-mutagenic based on the presence of experimental value in two models (CAESAR and SARpy/IRFMN) even though one of the models outcome does not agree with the other. The VEGA models whenever the experimental value is available for a molecule, give the experimental value as an output of the prediction, while T.E.S.T. presents the experimental value along with the results of the predictions. Contrarily, T.E.S.T. overall estimation for Diclofenac is mutagenic. The compound is not present in the training set either in the external test set of T.E.S.T. The consensus overall result of T.E.S.T. is the result of the integration of the predictions of the three models predictions (Hierarchical, FDA and Nearest neighbour). The Hierarchical model predicts the substance as non-mutagenic, while FDA and Nearest neighbour predictions are mutagenic. Together with these predictions also two tables of similar compounds extracted from the training and test sets are provided by T.E.S.T. Table 18 (page 113) reports the first two similar chemicals obtained from this table of similar compounds, with their similarity coefficient, experimental and predicted values. We selected the

first chemical similar to Diclofenac: N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS number: 66789-80-8), to further analyse the similarities and the differences between this chemical and the target chemical making use of ToxDelta.

Table 17. Summary of the prediction results of non-testing models for Diclofenac

Software	Model/method	Experimental value	Results	Predicted value	Applicability Domain Index
VEGA	CONSENSUS		Non-mutagenic		
	CAESAR	Non-mutagenic	Non-mutagenic		1
	SARpy/IRFMN	Non-mutagenic	Non-mutagenic		1
	ISS	N/A*	Non-mutagenic		0.772
	KNN/Read-across	N/A	Mutagenic		0.652
T.E.S.T.*	Consensus		Mutagenic	0.53	Internally checked
	Hierarchical	N/A	Non-Mutagenic	0.40	Internally checked
	FDA	N/A	Mutagenic	0.51	Internally checked
	Nearest neighbour	N/A	Mutagenic	0.67	Internally checked
ToxRead	Read-across		4 SAs:Non-mutagenic 1 SA: mutagenic		N/A

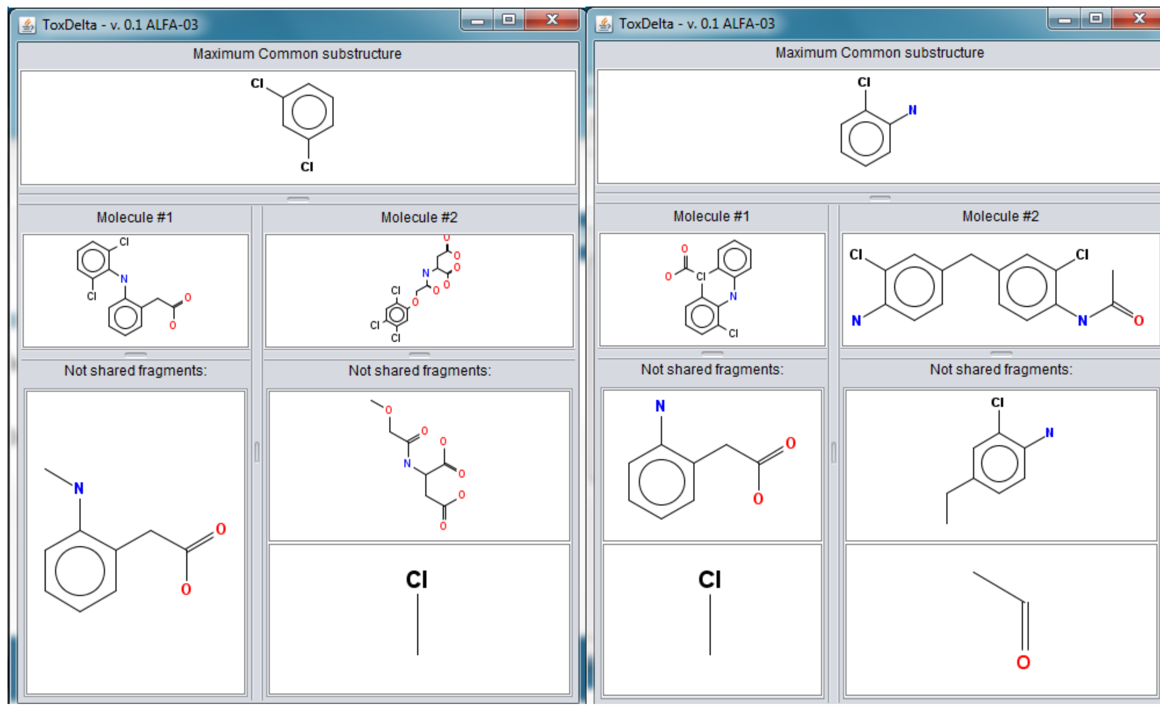
*Not availbale

Table 18. Experimental and prediction values for N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS 66789-80-8) and N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9) as examples of similar chemicals to Diclofenac in the outcome of T.E.S.T.

CAS	Structure	Similarity Coefficient	Experimental value	Predicted value
66789-80-8		0.61	0.00	0.02
91575-28-9		0.70	1.00	0.80

As shown in Table A.4 N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS 66789-80-8) and N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9) are the most similar chemicals identified by T.E.S.T (similarity coefficient: 0.61 and 0.70, respectively). While the experimental value and the predicted value of the first molecule are non-mutagenic, mutagenicity effect of the second molecule is indicated as positive. ToxDelta is used to investigate the structural similarities and dissimilarities, and the presence of mutagenic or non-mutagenic rules in Diclofenac and each similar substance. Figure 13 (page 115) shows the results of ToxDelta for the comparison between the target and each similar molecule. The MCS extracted from the two structures is illustrated on the top of the panel, while the dissimilar substructures of each molecule are listed below the corresponding molecules. In the case of the mutagenic similar compound (CAS 91575-28-9), on the right part of the figure, there are four fragments not in common between the target and the similar compound, two fragments present in the target, and two in the similar compound. The two fragments related to Diclofenac appear also as fragments in the comparison with the first similar compound, and we know that both Diclofenac and the similar compound are not toxic. Thus, we focus our attention to the other two fragments. The fragment CH3-C=O is not associated with any mutagenic activity. This can be verified for instance with ToxRead, studying Acetone. Figure 14 (page 116) reports the

dissimilar substructure in the target and the second similar chemical (CAS 91575-28-9) with mutagenic result. The presence of p-toluidine may trigger mutagenicity in a chemical. We also notice that the fragment Diphenylamine present in Diclofenac is a non-mutagenic SA.



(a) (b)

Figure 13. ToxDelta outcome for the comparison between Diclofenac and a) N-[(2,4,5-Trichlorophenoxy)acetyl]-L-aspartic acid (CAS 66789-80-8), b) N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9)

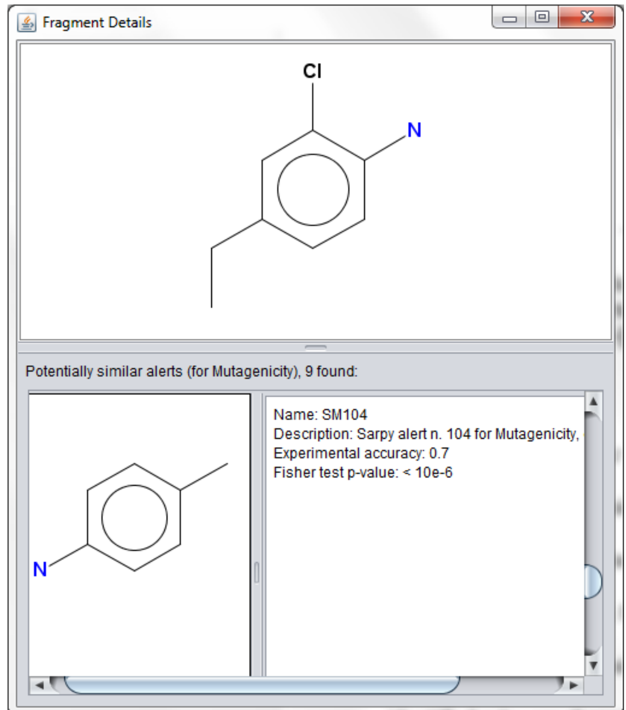


Figure 14. ToxDelta outcome of the dissimilar substructure remaining after the subtraction of the MCS and the structural rules identified inside the dissimilar substructure of N-[4-(4-Amino-3-chlorobenzyl)-2-chlorophenyl]acetamide (CAS 91575-28-9) (molecule #2)

Additionally, Diclofenac is analyzed by ToxRead to identify all the SAs present in the molecular structure, either mutagenic or non-mutagenic, together with the most structurally similar substances present in the database of ToxRead. Figure 15 (page 117) shows the outcome chart of ToxRead with the target substance in the middle, encircled by three similar substances (two mutagenic and one non-mutagenic). In addition, six SAs are identified and are connected to the target by straight arrows. These SAs consist of five non-mutagenic rules and one mutagenic rule. The list of the identified rules is reported in Table 19 (page 118). In fact, the mutagenic rule identified by ToxRead (MNM37) is a substructure of the non-mutagenic rule (SM197) present in the target substance, similar to the mutagenic rule identified by ToxDelta (Figure 16-page 118). The two mutagenic similar compounds indicated by ToxRead are analysed by ToxDelta for investigating the dissimilarities (Figure 17-page 118). Both similar chemicals contain Phenylamine SA as a mutagenic rule. As implied in the preceding section, this active SA is a part of a bigger SA identified in Diclofenac with non-mutagenic effect. This means the effect of the smaller SA is overcome by the parent SA which has a larger overlap

with the target compound. This important difference between the target and the similar compounds explains the different mutagenicity effect by means of a structure-based method, analysing the active and inactive rules inside the structure of the chemicals. Table 20 (page 119) is the tabular format for summarizing WOE assessment of Diclofenac in a qualitative way.

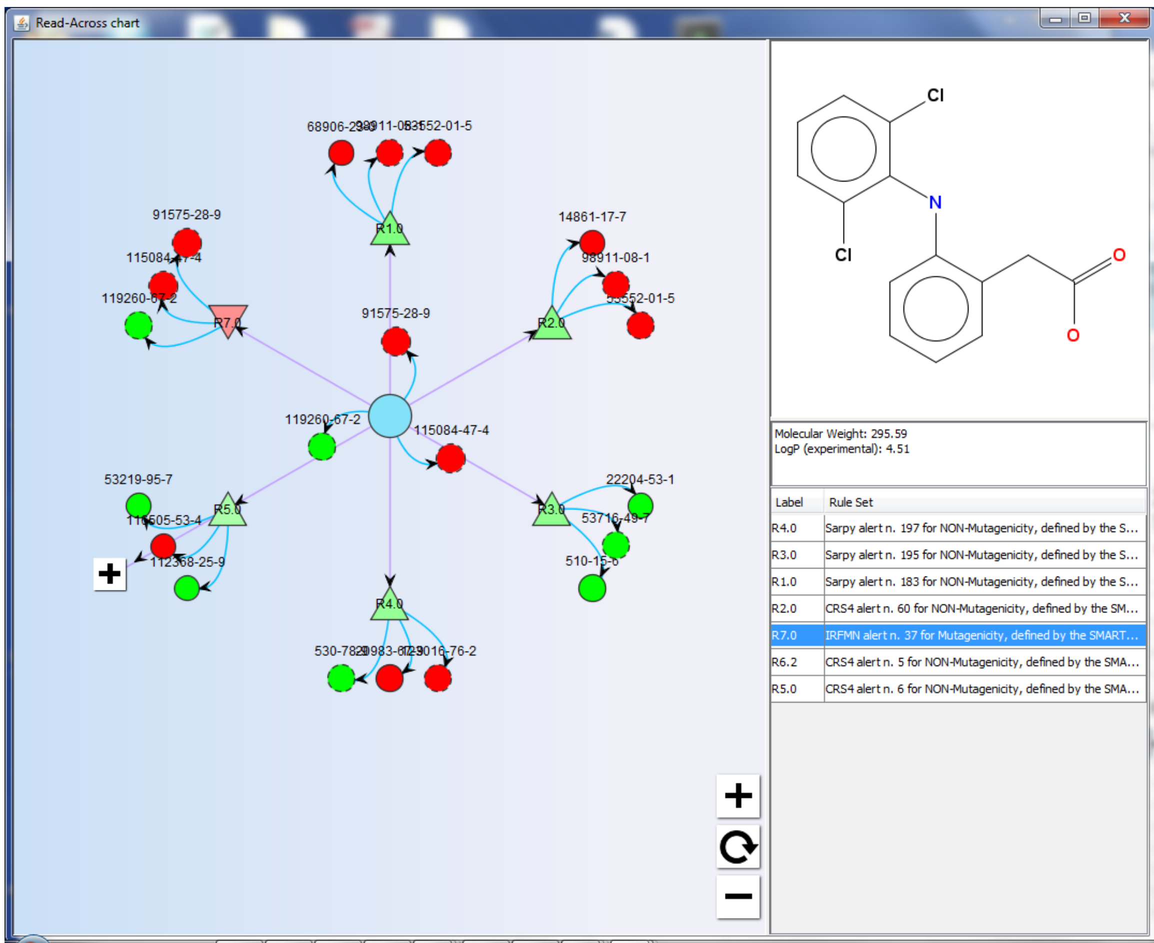


Figure 15. ToxRead chart for the target molecule Diclofenac. The target substance molecule is shown on the right side and the identified non-mutagenic and mutagenic rules are listed below

Table 19. The mutagenicity and non-mutagenicity rules identified by ToxRead in the structure of Diclofenac

Name	Effect	Source	Experimental accuracy	SMARTS	P-value
SM183	Non-mutagenic	SARpy	0.75	c1(cc(ccc1)Cl)Cl	<10e-6
CRM6	Non-mutagenic	CRS4	0.55	CC(=O)O	0.00075
CRM60	Non-mutagenic	CRS4	0.72	c1cc(cc(c1)Cl)Cl	0.00009
SM195	Non-mutagenic	SARpy	0.68	C(=O)(O)Cc1ccccc1	0.00511
SM197	Non-mutagenic	SARpy	0.65	N(c1ccccc1)c2ccccc2	0.00477
MNM37	Mutagenic	IRFMN	0.68	N- Ar	<10e-6

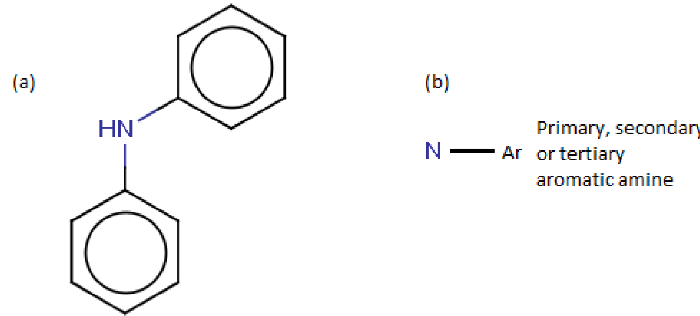


Figure 16. (a) SM197, Diphenylamine: non-mutagenic rule, (b) MNM37, Phenylamine: mutagenic rule identified by ToxRead

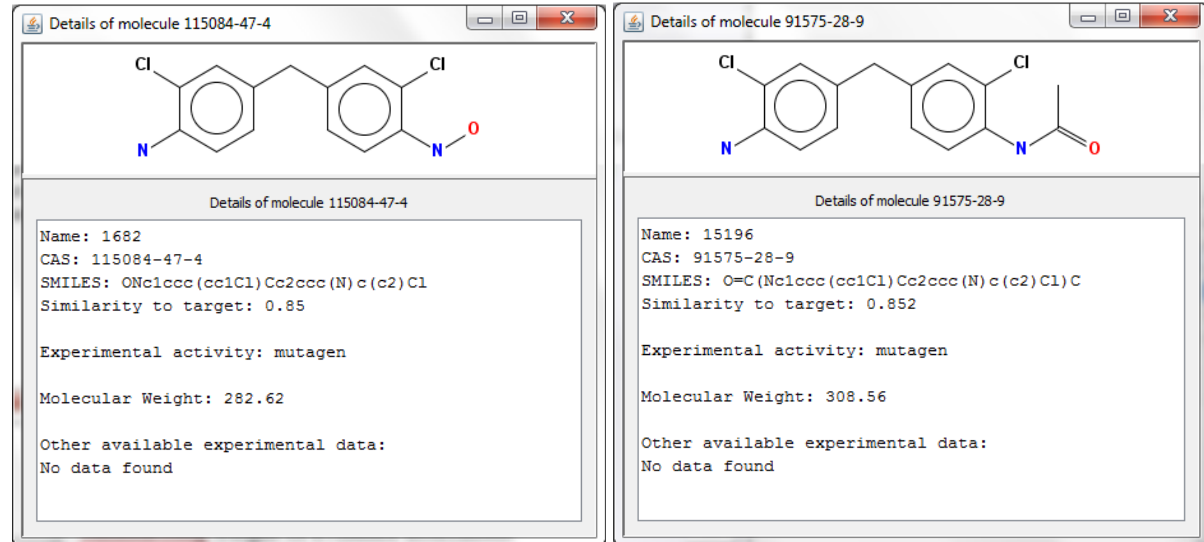


Figure 17. The two similar chemicals to Diclofenac extracted by ToxRead with mutagenic effect

Table 20. Optional tabular format for summarizing weight of evidence assessment of Diclofenac

Question		Hazard identification
Assemble the evidence	Select evidence	Nine (Q)SAR models from two *in silico* platforms, a read-across tool, ToxRead and a tool to investigate the dissimilarities between the similar compounds and the target, ToxDelta, are chosen for testing mutagenicity effect of Diclofenac. The used platforms provide both predicted values from several models, and indicate similar substances with experimental values, which can be used for read-across.
	Lines of evidence	The consensus model of VEGA estimated the chemical as non-mutagen. All the models in VEGA except KNN predicted the target as non-mutagenic. The consensus model of T.E.S.T. predicted the target as mutagenic. All the models in T.E.S.T. except Hierarchical predicted the compound as mutagenic with moderate predicted values. A number of similar compounds are suggested by T.E.S.T. The first two similar compounds are selected for a further investigation. The dissimilarities between the target and each similar compound is evaluated. Further analysis by ToxDelta, shows the presence of phenylamine in the similar mutagenic chemical which trigger the mutagenic effect. Although Phenylamine as a mutagenic SA with accuracy=0.68 exist in Diclofenac, it is a part of a bigger SA (Diphenylamine) with non-mutagenic effect. Due to this difference, the similarity of this chemical is not relevant to be used in a read-across way. Further evaluation by ToxRead, indicates five non-mutagenic rules and one mutagenic rule in the structure of the target. Similarly, in the results of ToxRead the mutagenic SA is Phenylamine which is present in the bigger non-mutagenic SA, Diphenyamine.
Weight the evidence	Methods	VEGA provides ADI that is a sort of quantitative measurement of reliability and values higher than 0.8 are considered more reliable. T.E.S.T. applies a filter to eliminate not reliable predictions. ToxRead indicates the structural alerts found in the target which are associated with the effect and an arbitrary number of similar compounds which share that rule with the target. ToxRead also reveals the SAs present in the source compounds, and by using ToxDelta user can check if these active moieties belong to the dissimilarities between the two molecules. For Diclofenac a further evaluation is performed on the similar compound indicated by T.E.S.T. and ToxRead to check the relevance and the reliability of the

		similar compounds to be used in terms of read-across.
	Results	All the predictions obtained from VEGA and T.E.S.T. are reliable in terms of ADI. A further evaluation is performed on the similar compound suggested by T.E.S.T. and ToxRead to check the relevance of the similarity to be used in read-across terms. Five non-mutagenic and one mutagenic rules in the structure of Diclofenac are identified by ToxRead. The results of dissimilarities indicated the presence of a mutagenic SA in Diclofenac which is a substructure of a bigger non-mutagenic SA, and therefore its toxic effect is not relevant.
Integrate the evidence	Methods	Inconsistency is present in the results of the two main platforms. ToxDelta and ToxRead both indicate the presence of a mutagenic rule in the structure of Diclofenac, but VEGA estimates the compound as non-mutagenic based on two experimental values present in its models. High ADI are assigned to individual predictions of the models, the conflict in the results is assessed by ToxDelta. Further investigation on the similar compounds suggested by T.E.S.T. and ToxRead is performed to evaluate the similarities and the differences between each pair of molecules and to reach a conclusion based on read-across. The reliability and relevance of all the similar compounds can be evaluated in this way. The non-relevance of phenylamine SA in the target makes the process of decision-making easier. Also the expert may find confidence in experimental results (in this case non-mutagenic effect).
	Results	The *in silico* methods used in this practice are not in concordance. The similar compounds suggested are useful for defining the reliability and relevance. Considering the evidence obtained from these predictions, it can be concluded that Diclofenac is not mutagenic.

References

1. Occupational Safety and Health Administration's (OSHA). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). (2013).
2. Benigni, R., Bossa, C. Structure alerts for carcinogenicity, and the Salmonella assay system: a novel insight through the chemical relational databases technology. *Mutation Research/Reviews in Mutation Research* **659,** 248–261 (2008).
3. Klaunig, J. E., Wang, Z., Pu, X., Zhou, S. Oxidative stress and oxidative damage in chemical carcinogenesis. *Toxicology and Applied Pharmacology* **254,** 86–99 (2011).
4. Pulliero, A. Godschalk, R., Andreassi, M. G., Curfs, D., Van Schooten, F. J., Izzotti, A. Environmental carcinogens and mutational pathways in atherosclerosis. *International Journal of Hygiene and Environmental Health* **218,** 293–312 (2015).
5. Tanaka, T., Shimizu, M., Kochi, T., Moriwaki, H. Chemical-induced carcinogenesis. *Journal of Experimental & Clinical Medicine* **5,** 203–209 (2013).
6. Vasseur, P., Lasne, C. OECD detailed review paper number 31 on "Cell transformation assays for detection of chemical carcinogens": Main results and conclusions. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **744,** 8–11 (2012).
7. Griffiths, A. J. F., Miller, J. H., Suzuki, D. T., Lewontin, R. C., Gelbart, W. M. An introduction to genetic analysis 7th edition, New York: W. H. *Freeman* (2000).
8. Lerman, L. S. The structure of the DNA-acridine complex. *Proceedings of the National Academy of Sciences* **49,** 94–102 (1963).
9. Conolly, R. B., Reitz, R. H., Clewell III, H. J., Andersen, M. E. Pharmacokinetics, biochemical mechanism and mutation accumulation: a comprehensive model of chemical carcinogenesis. *Toxicology Letters* **43**, 189-200 (1988).
10. van Leeuwen, I. M., Zonneveld, C. From exposure to effect: a comparison of modeling approaches to chemical carcinogenesis. *Mutation Research/Reviews in Mutation Research* **489**, 17-45 (2001).
11. Luch, A. Nature and nurture–lessons from chemical carcinogenesis. *Nature Reviews Cancer* **5**, 113 (2005).
12. Guengerich, F. P. Metabolic activation of carcinogens. *Pharmacology & Therapeutics* **54**, 17-61 (1992).
13. Park, B. K., Kitteringham, N. R., Maggs, J. L., Pirmohamed, M., Williams, D. P. The role of metabolic activation in drug-induced hepatotoxicity. *Annual Review of Pharmacology and Toxicology* **45**, 177-202 (2005).
14. Klaunig, J. E., Kamendulis, L. M., Xu, Y. Epigenetic mechanisms of chemical carcinogenesis. *Human & Experimental Toxicology* **19**, 543-555 (2000).
15. Ohshima, H., Tazawa, H., Sylla, B. S., Sawa, T. Prevention of human cancer by modulation of chronic inflammatory processes. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **591**, 110-122 (2005).
16. Galati, G., Teng, S., Moridani, M. Y., Chan, T. S., O'Brien, P. J. Cancer chemoprevention and apoptosis mechanisms induced by dietary polyphenolics. *Drug Metabolism and Drug Interactions* **17**, 311-350 (2000).
17. Shacter, E., Weitzman, S. A. Chronic inflammation and cancer. ONCOLOGY-WILLISTON PARK THEN HUNTINGTON- **16**, 217-229 (2002).
18. Guengerich, F. P. Metabolism of chemical carcinogens. *Carcinogenesis* **21**, 345-351 (2000).
19. Gonzalez, F. J. The use of gene knockout mice to unravel the mechanisms of toxicity and chemical carcinogenesis. *Toxicology Letters* **120**, 199-208 (2001).

20. Faccini, J. M., Butler, W. R., Friedmann, J. C., Hess, R., Reznik, G. K., Ito, N., *et al.* IFSTP guidelines for the design and interpretation of the chronic rodent carcinogenicity bioassay. *Experimental and Toxicologic Pathology* **44,** 443–456 (1992).
21. Maurici, D., Aardema, M., Corvi, R., Kleber, M., Krul, C. A. M., Laurent, C., *et al.* Genotoxicty and mutagenicity. *ATLA Alternatives to Laboratory Animals* **Suppl 1, 33,** 117–130 (2005).
22. UNECE. UNECE Revised manual on methodologies and criteria for mapping critical levels/loads and geographical areas where they are exceeded. (2004).
23. Cohen, S. M., Ellwein, L. B. Genetic errors, cell proliferation, and carcinogenesis. *Cancer Research* **51,** 6493–6505 (1991).
24. Weinberg, R. A. Oncogenes, antioncogenes, and the molecular bases of multistep carcinogenesis. *Cancer Research* **49,** 3713–3721 (1989).
25. McDonnell, T. J. Cell division versus cell death: a functional model of multistep neoplasia. *Molecular Carcinogenesis* **8,** 209–213 (1993).
26. Page, N. Concept of a bioassay program in environmental carcinogenesis. *In:* Environmental Cancer: Advances in Modern Toxicology. New York, *John Wiley and Sons* (1977).
27. Cohen, S. M., Cano, M., Earl, R. A., Carson, S. D., Garland, E. M. A proposed role for silicates and protein in the proliferative effects of saccharin on the male rat urothelium. *Carcinogenesis* **12,** 1551–1555 (1991).
28. Monro, A. What is an appropriate measure of exposure when testing drugs for carcinogenicity in rodents? *Toxicology and Applied Pharmacology* **112,** 171–181 (1992).
29. Vater, S. T., McGinnis, P. M., Schoeny, R. S., Velazquez, S. F. Biological considerations for combining carcinogenicity data for quantitative risk assessment. *Regulatory Toxicology and Pharmacology* **18,** 403–418 (1993).
30. Goodman, G., Wilson, R. Predicting the carcinogenicity of chemicals in humans from rodent bioassay data. *Environmental Health Perspectives* **94,** 195 (1991).
31. Vainio, H., Heseltine, E., McGregor, D., Tomatis, L., Wilbourn, J. Working group on mechanisms of carcinogenesis and the evaluation of carcinogenic risks. 2357-2361 (1992).
32. Melnick, R. L., Kohn, M. C., Portier, C. J. Implications for risk assessment of suggested nongenotoxic mechanisms of chemical carcinogenesis. *Environmental Health Perspectives* **104,** 123 (1996).
33. Williams, G. M. Mechanisms of chemical carcinogenesis and application to human cancer risk assessment. *Toxicology* **166,** 3–10 (2001).
34. European Parliament No. 1272/2008 of the European Parliament and of the Council, on Classification, Labeling and Packaging of Substances and Mixtures, Amending and Repealing Directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No. 1907/2006, Official J. *Eur. Union* L353 (2008).
35. Goetz, M. E., Luch, A. Reactive species: a cell damaging rout assisting to chemical carcinogens. *Cancer Letters* **266,** 73–83 (2008).
36. Enoch, S. J., Cronin, M. T. D. Development of new structural alerts suitable for chemical category formation for assigning covalent and non-covalent mechanisms relevant to DNA binding. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **743,** 10–19 (2012).
37. Nash, H. M., Lu, R., Lane, W. S., Verdinel, G. L. The critical active-site amine of the human 8-oxoguanine DNA glycosylase, hOgg1: direct identification, ablation and chemical reconstitution. *Chemistry & Biology* **4,** 693–702 (1997).
38. Enoch, S. J., Cronin, M. T., Ellison, C. M. The use of a chemistry-based profiler for covalent DNA binding in the development of chemical categories for read-across for genotoxicity. *Alternatives to Laboratory Animals: ATLA* **39,** 131–145 (2011).
39. Benigni, R., Bossa, C., Tcheremenskaia, O. Nongenotoxic carcinogenicity of chemicals: mechanisms of action and early recognition through a new set of structural alerts. *Chemical Reviews* **113,** 2940–2957 (2013).
40. Kulis, M., and Esteller, M. DNA methylation and cancer. *In*: Advances in genetics. Vol. **70.** *Academic Press* 27-56 (2010).

41. Ropero, S., Esteller, M. The role of histone deacetylases (HDACs) in human cancer. *Molecular oncology* **1,** 19–25 (2007).

42. Zhong, M., Nie, X., Yan, A., Yuan, Q. Carcinogenicity prediction of noncongeneric chemicals by a support vector machine. *Chemical Research in Toxicology* **26,** 741–749 (2013).

43. Benigni, R. The Benigni/Bossa rulebase for mutagenicity and carcinogenicity–a module of Toxtree. *JRC Scientific and Technical Reports* 1–78 (2008).

44. Snodin, D. J. Genotoxic impurities: from structural alerts to qualification. *Organic Process Research & Development* **14,** 960–976 (2010).

45. Ellison, C. M., Sherhod, R., Cronin, M. T., Enoch, S. J., Madden, J. C., Judson, P. N. Assessment of methods to define the applicability domain of structural alert models. *Journal of Chemical Information and Modeling* **51,** 975–985 (2011).

46. Kazius, J., McGuire, R., Bursi, R. Derivation and validation of toxicophores for mutagenicity prediction. *Journal of Medicinal Chemistry* **48,** 312–320 (2005).

47. Miller, J. Miller, E. Origins of human cancer. *Cold Spring Harbor Laboratory, Cold Spring Harbor, NY* 605–627 (1977).

48. Ashby, J., Tennant, R. W. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. *Mutation Research/Genetic Toxicology* **204,** 17–115 (1988).

49. Munro, I. C., Ford, R. A., Kennepohl, E., Sprenger, J. G. Thresholds of toxicological concern based on structure-activity relationships. *Drug Metabolism Reviews* **28,** 209–217 (1996).

50. Cheeseman, M. A., Machuga, E. J., Bailey, A. B. A tiered approach to threshold of regulation. *Food and Chemical Toxicology* **37,** 387–412 (1999).

51. Kroes, R., Renwick, A. G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., *et al.* Structure-based thresholds of toxicological concern (TTC): guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology* **42,** 65–83 (2004).

52. Patlewicz, G., Jeliazkova, N., Safford, R. J., Worth, A. P., Aleksiev, B. An evaluation of the implementation of the Cramer classification scheme in the Toxtree software. *SAR and QSAR in Environmental Research*, **19**, 495-524 (2008). https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/qsar_tools/toxtree (Accessed 03 May 2018).

53. National Research Council. Use of the maximum tolerated dose in animal bioassays for carcinogenicity. *Issues in Risk Assessment* 15-83 (1993).

54. Clayson, D. B., Clegg, D. J. Classification of carcinogens: polemics, pedantics, or progress? *Regulatory Toxicology and Pharmacology* **14,** 147–166 (1991).

55. Weisburger, J. H., Williams, G. M. Bioassay of carcinogens: in vitro and in vivo tests. *Chemical Carcinogens* **2,** 1323–1373 (1984).

56. NAVIGATOR, REACH. Guidance for the Implementation of REACH–Guidance on Information Requirements and Chemical Safety Assessment, *Part E: Risk Characterization, European Chemicals Agency* (2008).

57. European Chemical Agency (ECHA). Evaluation under REACH, progress report 2013. European Chemicals Bureau (2013). https://echa.europa.eu/documents/10162/13628/evaluation_report_2013_en.pdf (Accessed 03 May 2018).

58. Worth, A. P., Bassan, A., De Bruijn, J., Gallegos Saliner, A., Netzeva, T., Pavan, M., *et al.* The role of the European Chemicals Bureau in promoting the regulatory use of (Q) SAR methods. *SAR and QSAR in Environmental Research* **18,** 111–125 (2007).

59. European Chemical Agency (ECHA). Guidance on information requirements and chemical safety assessment. Chapter R 8**,** (2008).

60. Scialli, A. R. The challenge of reproductive and developmental toxicology under REACH. *Regulatory Toxicology and Pharmacology* **51,** 244–250 (2008).

61. REACH - Registration, Evaluation, Authorisation and Restriction of Chemicals. [Annex VI] Available at: http://www.reachonline.eu/REACH/EN/REACH_EN/articleVI.html (Accessed 03 May 2018).

62. Council, E. E. C. EEC Council Directive 86/609/EEC of 24 November 1986 on the approximation of laws, regulations and administrative provisions of the Member States regarding the protection of animals used for experimental and other scientific purposes. *Official Journal of European Union L* **358,** 1–28 (1986).

63. Corvi, R., Madia, F., Worth, A., Whelan, M. EURL ECVAM strategy to avoid and reduce animal use in genotoxicity testing. *JRC Scientific and Policy Reports* 48 (2013).

64. Adler, S., Basketter, D., Creton, S., Pelkonen, O., Van Benthem, J., Zuang, V., *et al.* Alternative (non-animal) methods for cosmetics testing: current status and future prospects-2010. *Archives of Toxicology* **85,** 367–485 (2011).

65. Pfuhler, S., Kirkland, D., Kasper, P., Hayashi, M., Vanparys, P., Carmichael, P., *et al.* Reduction of use of animals in regulatory genotoxicity testing: identification and implementation opportunities—report from an ECVAM workshop. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **680,** 31–42 (2009).

66. ECETOC. Category approaches, Read- across, (Q)SAR. Technical Report No. http://www.ecetoc.org/wp-content/uploads/2014/08/ECETOC-TR-116-Category-approaches-Read-across-QSAR.pdf (Accessed 03 May 2018).

67. OECD. OECD guideline for the testing of chemicals. Test No. 453: Combined Chronic Toxicity/Carcinogenicity Studies. *OECD Publishing* (2009). https://read.oecd-ilibrary.org/environment/test-no-453-combined-chronic-toxicity-carcinogenicity-studies_9789264071223-en#page1 (Accessed 03 May 2018).

68. Robinson, D. E., Macdonald, J. S. Background and framework for ILSI's collaborative evaluation program on alternative models for carcinogenicity assessment. *Toxicologic Pathology* **29,** 13–19 (2001).

69. Goodman, J. I. A perspective on current and future uses of alternative models for carcinogenicity testing. *Toxicologic Pathology* **29,** 173–176 (2001).

70. Rossello, R. A., Kohn, D. H. Gap junction intercellular communication: a review of a potential platform to modulate craniofacial tissue engineering. *Journal of Biomedical Materials Research Part B: Applied Biomaterials* **88,** 509–518 (2009).

71. Combes, R. D. The use of structure–activity relationships and markers of cell toxicity to detect non-genotoxic carcinogens. *Toxicology in Vitro* **14,** 387–399 (2000).

72. OECD Environment Directorate, Environment, Health & and Safety Division. OECD Environment Health and Safety Publications Series on Testing and Assessment No. 49. The report from the expert group on (quantitative) structure-activity relationships [(Q)SARS] on the principles for the validation of (Q)SARs 2nd Meeting of the ad hoc Expert Group on QSARs OECD Headquarters, 20-21 (2004).

73. OECD. OECD Principles for the validation, for regulatory purposes, of (quantitative) structure-activity relationship models. (2004).

74. Uno, Y. *et al.* Rat liver in vivo replicative DNA synthesis test for short-term prediction of nongenotoxic (Ames-negative) hepatocarcinogenicity: a collaborative study of the Nongenotoxic Carcinogen Study Group of Japan. *Toxicology Letters* **109,** 105–114 (1999).

75. Uno, Y., Matsuura, K., Miyagawa, M., Takasawa, H., Tanifuji, H., Abe, K., *et al.* An in vivo-in vitro replicative DNA synthesis (RDS) test using rat hepatocytes as an early prediction assay for nongenotoxic hepatocarcinogens screening of 22 known positives and 25 noncarcinogens. *Mutation Research/Genetic Toxicology* **320,** 189–205 (1994).

76. Yoshifumi, U., Hironao, T., Makoto, M., Yuki, I., Taeko, M., Masuo, O., *et al.* In vivo-in vitro replicative DNA synthesis (RDS) test using perfused rat livers as an early prediction assay for nongenotoxic hepatocarcinogens: I. Establishment of a standard protocol. *Toxicology Letters* **63,** 191–199 (1992).

77. Yoshikawa, K. Anomalous nonidentity between Salmonella genotoxicants and rodent carcinogens: nongenotoxic carcinogens and genotoxic noncarcinogens. *Environmental Health Perspectives* **104,** 40 (1996).

78. Golbamaki, A., Benfenati, E., Golbamaki, N., Manganaro, A., Merdivan, E., Roncaglioni, A., Gini, G., *et al.* New clues on carcinogenicity-related substructures derived from mining two large datasets of chemical compounds. *Journal of Environmental Science and Health, Part C* **34,** 97–113 (2016).

79. Sanderson, D., Earnshaw, C. Computer prediction of possible toxic action from chemical structure; the DEREK system. *Human & Experimental Toxicology* **10,** 261–273 (1991).

80. Judson, P. N., Stalford, S. A., Vessey, J. Assessing confidence in predictions made by knowledge-based systems. *Toxicology Research* **2,** 70–79 (2013).

81. Williams, R. V., Amberg, A., Brigo, A., Coquin, L., Giddings, A., Glowienke, S., *et al.* It's difficult, but important, to make negative predictions. *Regulatory Toxicology and Pharmacology* **76,** 79–86 (2016).

82. TOPKAT, version 3. User Guide, Accelrys software Inc. San Diego, CA, USA. http://accelrys.com. (Accessed 03 May 2018).

83. Klopman, G. MULTICASE 1. A hierarchical computer automated structure evaluation program. *Molecular Informatics* **11**, 176-184. (1992). http://www.multicase.com/. (Accessed 03 May 2018).

84. OECD QSAR Toolbox - OECD. http://www.oecd.org/chemicalsafety/riskassessment/theoecdqsartoolbox.htm (Accessed 03 May 2018).

85. Helma, C. Lazy structure-activity relationships (lazar) for the prediction of rodent carcinogenicity and Salmonella mutagenicity. *Molecular Diversity* **10,** 147–158 (2006).

86. ACD/Tox Suite. Version 2.9, Data Sheet, Advanced Chemistry Development, Inc. Toronto, On, Canada, (2012) http://www.acdlabs.com (Accessed 03 May 2018).

87. Roberts, G., Myatt, G. J., Johnson, W. P., Cross, K. P., Blower, P. E. LeadScope: software for exploring large sets of screening data. *Journal of Chemical Information and Computer Sciences* **40**, 1302-1314. (2000). http://www.leadscope.com/ (Accessed 03 May 2018).

88. SARpy. Version 1.0 http://sarpy.sourceforge.net/ (Accessed 03 May 2018).

89. Ferrari, T., Cattaneo, D., Gini, G., Golbamaki Bakhtyari, N., Manganaro, A., Benfenati, E. Automatic knowledge extraction from chemical structures: the case of mutagenicity prediction. *SAR and QSAR in Environmental Research* **24,** 365–383 (2013).

90. Chapman, P. M. Presentation and interpretation of sediment quality triad data. *Ecotoxicology* **5,** 327–339 (1996).

91. Hardy, A., Benford, D., Halldorsson, T., Jeger, M. J., Knutsen, H. K., More, S., *et al.* Guidance on the use of the weight of evidence approach in scientific assessments. *EFSA Journal* **15,** (2017).

92. Manganelli, S., Schilter, B., Benfenati, E., Manganaro, A., Piparo, E. L. Integrated strategy for mutagenicity prediction applied to food contact chemicals. *ALTEX-Alternatives to Animal Experimentation* **35**, 169-178. (2018).

93. Smith, E. P., Lipkovich, I., Ye, K. Weight-of-evidence (WOE): quantitative estimation of probability of impairment for individual and multiple lines of evidence. *Human and Ecological Risk Assessment* **8,** 1585–1596 (2002).

94. Linkov, I., Satterstrom, F. K., Kiker, G., Batchelor, C., Bridges, T., Ferguson, E. From comparative risk assessment to multi-criteria decision analysis and adaptive management: Recent developments and applications. *Environment International* **32,** 1072–1093 (2006).

95. Zuin, S., Micheletti, C., Critto, A., Pojana, G., Johnston, H., Stone, V., *et al.* Weight of evidence approach for the relative hazard ranking of nanomaterials. *Nanotoxicology* **5,** 445–458 (2011).

96. Linkov, I., Satterstrom, F. K., Steevens, J., Ferguson, E., Pleus, R. C. Multi-criteria decision analysis and environmental risk assessment for nanomaterials. *Journal of Nanoparticle Research* **9,** 543–554 (2007).

97. Linkov, I., Satterstrom, F. K., Corey, L. M. Nanotoxicology and nanomedicine: making hard decisions. *Nanomedicine: Nanotechnology, Biology and Medicine* **4,** 167–171 (2008).

98. Linkov, I., Loney, D., Cormier, S., Satterstrom, F. K., Bridges, T. Weight-of-evidence evaluation in environmental assessment: review of qualitative and quantitative approaches. *Science of the Total Environment* **407,** 5199–5205 (2009).

99. Weed, D. L. Weight of evidence: a review of concept and methods. *Risk Analysis* **25,** 1545–1557 (2005).

100. Chapman, P. M., McDonald, B. G., Lawrence, G. S. Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. *Human and Ecological Risk Assessment* **8,** 1489–1515 (2002).

101. USEPA. USEPA guidelines for carcinogen risk assessment. Risk Assessment Forum U.S. Environmental Protection Agency Washington, DC (2005). https://www3.epa.gov/airtoxics/cancer_guidelines_final_3-25-05.pdf (Accessed 03 May 2018).

102. Oberdorster, G., Oberdorster, E., Oberdorster, J. Nanotoxicology: an emerging discipline evolving from studies of ultrafine particles. *Environmental Health Perspectives* **113,** 823–839 (2005).

103. Thomas, K., Sayre, P. Research strategies for safety evaluation of nanomaterials, Part I: evaluating the human health implications of exposure to nanoscale materials. *Toxicological Sciences* **87,** 316–321 (2005).

104. Linkov, I., Satterstrom, F. K. Nanomaterial risk assessment and risk management. *In:* Real-Time and Deliberative Decision Making 129–157, *Springer* (2008).

105. Morgan, K. Development of a preliminary framework for informing the risk analysis and risk management of nanoparticles. *Risk Analysis* **25,** 1621–1635 (2005).

106. TGD, EU. Technical guidance document on risk assessment in support of commission directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on Risk Assessment for existing substances, and Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Part I–IV, European Chemicals Bureau (ECB), JRC-Ispra (VA), Italy, April 2003. *Part II. European Commission Joint Research Centre. EUR* **20418,** (2003).

107. ANTARES. http://www.antares-life.eu/ (Accessed 03 May 2018).

108. Kirkland, D., Kasper, P., Müller, L., Corvi, R., Speit, G. Recommended lists of genotoxic and non-genotoxic chemicals for assessment of the performance of new or improved genotoxicity tests: a follow-up to an ECVAM workshop. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis* **653,** 99–108 (2008).

109. ISSCAN database. http://old.iss.it/meca/index.php?lang=1&anno=2013&tipo=25 (Accessed 03 May 2018)

110. Martin, Y. C., Kofron, J. L., Traphagen, L. M. Do structurally similar molecules have similar biological activity? *Journal of Medicinal Chemistry* **45,** 4350–4358 (2002).

111. Kubinyi, H. Chemical similarity and biological activities. *Journal of the Brazilian Chemical Society* **13,** (2002).

112. Gold, L. S., Manley, N. B., Slone, T. H., Rohrbach, L., Garfinkel, G. B. Supplement to the Carcinogenic Potency Database (CPDB): results of animal bioassays published in the general literature through 1997 and by the National Toxicology Program in 1997-1998. *Toxicological Sciences* **85,** 747–808 (2005).

113. istMolBase. (Kode). https://chm.kode-solutions.net/products_istmolbase.php (Accessed 03 May 2018).

114. InstantJChem. Calculation module developed by ChemAxon. http://www.chemaxon.com/ (Accessed 03 May 2018).

115. Weininger, D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *Journal of Chemical Information and Computer Sciences* **28,** 31–36 (1988).

116. OECD eChemPortal. https://www.echemportal.org/echemportal/index.action (Accessed 03 May 2018).

117. European Chemicals Agency (ECHA) CLP inventory. https://echa.europa.eu/information-on-chemicals/cl-inventory-database (Accessed 03 May 2018).

118. SMARTS Theory. DayLight. Chemical information system Inc. http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html (Accessed 03 May 2018).

119. Ferrari, T., Gini, G., Bakhtyari, N. G., Benfenati, E. Mining toxicity structural alerts from SMILES: A new way to derive Structure Activity Relationships. *In*: Computational Intelligence and Data Mining (CIDM), *2011 IEEE Symposium on* 120-127 (2011).

120. Tropsha, A., Gramatica, P., Gombar, V. K. The importance of being earnest: validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR & Combinatorial Science* **22,** 69–77 (2003).

121. Perkins, R., Fang, H., Tong, W., Welsh, W. J. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. *Environmental Toxicology and Chemistry* **22,** 1666–1679 (2003).

122. Benfenati, E., Manganelli, S., Giordano, S., Raitano, G., Manganaro, A. Hierarchical Rules for Read-Across and In Silico Models of Mutagenicity. *Journal of Environmental Science and Health, Part C* **33,** 385–403 (2015).

123. Gini, G., Franchi, A. M., Manganaro, A., Golbamaki, A., Benfenati, E. ToxRead: a tool to assist in read across and its use to assess mutagenicity of chemicals. *SAR and QSAR in Environmental Research* **25,** 999–1011 (2014).

124. Cao, Y., Jiang, T., Girke, T. A maximum common substructure-based algorithm for searching and predicting drug-like compounds. *Bioinformatics* **24,** i366–i374 (2008).

125. Floris, M., Manganaro, A., Nicolotti, O., Medda, R., Mangiatordi, G. F., Benfenati, E. A generalizable definition of chemical similarity for read-across. *Journal of Cheminformatics* **6,** 39 (2014).

126. Vernon, R. E. Which Elements Are Metalloids? *Journal of Chemical Education* **90,** 1703–1707 (2013).

127. Landsiedel, R., Kapp, M. D., Schulz, M., Wiench, K., Oesch, F. Genotoxicity investigations on nanomaterials: methods, preparation and characterization of test material, potential artifacts and limitations—many questions, some answers. *Mutation Research/Reviews in Mutation Research* **681,** 241–258 (2009).

128. Golbamaki, N., Rasulev, B., Cassano, A., Robinson, R. L. M., Benfenati, E., Leszczynski, J., Cronin, M. T. Genotoxicity of metal oxide nanomaterials: Review of Recent data and discussion of possible mechanisms. *Nanoscale* (2014).

129. Lubinski, L., Urbaszek, P., Gajewicz, A., Cronin, M. T. D., Enoch, S. J., Madden, J. C., *et al.* Evaluation criteria for the quality of published experimental data on nanomaterials and their usefulness for QSAR modelling. *SAR* and *QSAR* in *Environmental Research* **24,** 995–1008 (2013).

130. Marchese Robinson, R. L., Cronin, M. T., Richarz, A. N., Rallo, R. An ISA-TAB-Nano based data collection framework to support data-driven modelling of nanotoxicology. *Beilstein Journal of Nanotechnology* **6,** 1978–1999 (2015).

131. Stefaniak, A. B., Hackley, V. A., Roebben, G., Ehara, K., Hankin, S., Postek, M. T., *et al.* Nanoscale reference materials for environmental, health and safety measurements: needs, gaps and opportunities. *Nanotoxicology* **7,** 1325–1337 (2013).

132. Singh, N. P., McCoy, M. T., Tice, R. R., Schneider, E. L. A simple technique for quantitation of low levels of DNA damage in individual cells. *Experimental cell research* **175,** 184–191 (1988).

133. Hristozov, D. R., Zabeo, A., Foran, C., Isigonis, P., Critto, A., Marcomini, A., Linkov, I. A weight of evidence approach for hazard screening of engineered nanomaterials. *Nanotoxicology* **8,** 72–87 (2014).

134. Sekar, D., Falcioni, M. L., Barucca, G., Falcioni, G. DNA damage and repair following In vitro exposure to two different forms of titanium dioxide nanoparticles on trout erythrocyte. *Environmental Toxicology* **29,** 117–127 (2014).

135. Kim, Y. J., Choi, H. S., Song, M. K., Youk, D. Y., Kim, J. H., Ryu, J. C. Genotoxicity of aluminium oxide (Al2O3) nanoparticle in mammalian cell lines. *Molecular & Cellular Toxicology* **5,** 172–178 (2009).

136. Demir, E., Burgucu, D., Turna, F., Aksakal, S., Kaya, B. Determination of TiO2, ZrO2, and Al2O3 Nanoparticles on Genotoxic Responses in Human Peripheral Blood Lymphocytes and Cultured Embyronic Kidney Cells. *Journal of Toxicology and Environmental Health, Part A* **76,** 990–1002 (2013).

137. Auffan, M., Decome, L., Rose, J., Orsiere, T., De Meo, M., Briois, V., *et al.* In vitro interactions between DMSA-coated maghemite nanoparticles and human fibroblasts: A physicochemical and cyto-genotoxical study. *Environmental Science and Technology* **40,** 4367–4373 (2006).

138. Guichard, Y., Schmit, J., Darne, C., Gaté, L., Goutet, M., Rousset, D., *et al.* Cytotoxicity and genotoxicity of nanosized and microsized titanium dioxide and iron oxide particles in Syrian hamster embryo cells. *Annals of Occupational Hygiene* **56,** 631–644 (2012).

139. Könczöl, M., Ebeling, S., Goldenberg, E., Treude, F., Gminski, R., Gieré, R., *et al.* Cytotoxicity and genotoxicity of size-fractionated iron oxide (magnetite) in A549 human lung epithelial cells: role of ROS, JNK, and NF-κB. *Chemical Research in Toxycology* **24,** (2011).

140. Demir, E., Kaya, N., KAYA, B. Genotoxic effects of zinc oxide and titanium dioxide nanoparticles on root meristem cells of Allium cepa by comet assay. *Turkish Journal of Biology* **38,** 31–39 (2014).

141. Gražulis, S., Daškevič, A., Merkys, A., Chateigner, D., Lutterotti, L., Quiros, M., *et al.* Crystallography Open Database (COD): an open-access collection of crystal structures and platform for world-wide collaboration. *Nucleic Acids Research* **40,** D420-D427 (2012).

142. Gajewicz, A., Puzyn, T., Rasulev, B., Leszczynska, D., Leszczynski, J. Metal oxide nanoparticles: size-dependence of quantum-mechanical properties. *Nanoscience and Nanotechnology - Asia* **1,** 53–58 (2011).

143. Dennington, R., Keith, T., Millam, J. GaussView Version 5. (2009). http://gaussian.com/gaussview6/ (Accessed 03 May 2018)

144. Stewart, J. J. P. *MOPAC*. Colorado Springs (2012). http://openmopac.net/ (Accessed 03 May 2018).

145. Puzyn, T., Suzuki, N., Haranczyk, M., Rak, J. Calculation of quantum-mechanical descriptors for QSPR at the DFT level: is it necessary? *Journal of Chemical Information and Modelling* **48,** 1174–1180 (2008).

146. Mason, C. H., Perreault Jr, W. D. Collinearity, power, and interpretation of multiple regression analysis. *Journal of Marketing Research* 268–280 (1991).

147. Kiers, H. A., Smilde, A. K. A comparison of various methods for multivariate regression with highly collinear variables. *Statistical Methods and Applications* **16,** 193–228 (2007).

148. Therneau, T. M., Atkinson, E. J. An introduction to recursive partitioning using the RPART routines. Technical report Mayo Foundation (1997).

149. Hansen, K., Mika, S., Schroeter, T., Sutter, A., Ter Laak, A., Steger-Hartmann, T., *et al.* Benchmark data set for in silico prediction of Ames mutagenicity. *Journal of Chemical Information and Modelling* **49,** 2077–2081 (2009).

150. Kukreja, L., Barik, S., Misra, P. Variable band gap ZnO nanostructures grown by pulsed laser deposition. *Journal of Crystal Growth* **268,** 531–535 (2004).

151. Qu, Z., Kroes, G.-J. Theoretical study of the electronic structure and stability of titanium dioxide clusters (TiO2) n with n= 1- 9. *The Journal of Physical Chemistry B* **110,** 8998–9007 (2006).

152. Zhai, H.-J., Wang, L.-S. Probing the Electronic Structure and Band Gap Evolution of Titanium Oxide Clusters (TiO2) n-(n= 1- 10) Using Photoelectron Spectroscopy. *Journal of the American Chemical Society* **129,** 3022–3026 (2007).

153. Adams, L. K., Lyon, D. Y., Alvarez, P. J. Comparative eco-toxicity of nanoscale TiO 2, SiO 2, and ZnO water suspensions. *Water Research* **40,** 3527–3532 (2006).

154. Huk, A., Collins, A. R., El Yamani, N., Porredon, C., Azqueta, A., de Lapuente, J., *et al.* Critical factors to be considered when testing nanomaterials for genotoxicity with the comet assay. *Mutagenesis* **30,** 85–88 (2015).

155. Tice, R. R., Agurell, E., Anderson, D., Burlinson, B., Hartmann, A., Kobayashi, H., *et al.* Single cell gel/comet assay: Guidelines for in vitro and in vivo genetic toxicology testing. *Environmental and Molecular Mutagenesis* **35,** 206–221 (2000).

156. Singh, N., Manshian, B., Jenkins, G. J., Griffiths, S. M., Williams, P. M., Maffeis, T. G., *et al.* NanoGenotoxicology: The DNA damaging potential of engineered nanomaterials. *Biomaterials* **30,** 3891–3914 (2009).

157. Collins, A. R. The comet assay for DNA damage and repair: Principles, applications, and limitations. *Applied Biochemistry and Biotechnology - Part B Molecular Biotechnology* **26,** 249–261 (2004).

158. Collins, A. R., El Yamani, N., Lorenzo, Y., Shaposhnikov, S., Brunborg, G., Azqueta, A. Controlling variation in the comet assay. *Frontiers in Genetics* **5,** (2014).

159. Dhawan, A., Bajpayee, M., Parmar, D. Comet assay: a reliable tool for the assessment of DNA damage in different models. *Cell Biology and Toxicology* **25,** 5–32 (2009).

160. Ersson, C., Møller, P., Forchhammer, L., Loft, S., Azqueta, A., Godschalk, R. W., *et al.* An ECVAG inter-laboratory validation study of the comet assay: inter-laboratory and intra-laboratory variations of DNA strand breaks and FPG-sensitive sites in human mononuclear cells. *Mutagenesis* **28,** 279–286 (2013).

161. Forchhammer, L., Johansson, C., Loft, S., Möller, L., Godschalk, R. W., Langie, S. A., *et al.* Variation in the measurement of DNA damage by comet assay measured by the ECVAG inter-laboratory validation trial. *Mutagenesis* **25,** 113–123 (2010).

162. Forchhammer, L., Ersson, C., Loft, S., Möller, L., Godschalk, R. W., Van Schooten, F. J., *et al.* Inter-laboratory variation in DNA damage using a standard comet assay protocol. *Mutagenesis* **27,** 665–672 (2012).

163. Johansson, C., Møller, P., Forchhammer, L., Loft, S., Godschalk, R. W., Langie, S. A., *et al.* An ECVAG trial on assessment of oxidative damage to DNA measured by the comet assay. *Mutagenesis* **25,** 125–132 (2010).

164. Puzyn, T., Rasulev, B., Gajewicz, A., Hu, X., Dasari, T. P., Michalkova, A., *et al.* Using nano-QSAR to predict the cytotoxicity of metal oxide nanoparticles. *Nature Nanotechnology* **6,** 175–178 (2011).

165. Fu, J., Rong, G., Deng, Y. Mammalian cell cytotoxicity and genotoxicity of metallic nanoparticles. *Advanced Science Letters* **5,** 294–298 (2012).

166. Turkez, H., Celik, K., Cakmak, B. Biosafety evaluation of nanoparticles in view of genotoxicity and carcinogenicity studies: A systematic review. *In*: Key Engineering Materials, *Trans Tech Publications.* **543,** 200–203 (2013).

167. VEGA. http://www.vega-qsar.eu/ (Accessed 03 May 2018).

168. FDA. U. S. Food and Drug Administration (FDA) Genetic Toxicity, Reproductive and Development Toxicity, and Carcinogenicity Database (2009).

169. Magdolenova, Z., Collins, A., Kumar, A., Dhawan, A., Stone, V., Dusinska, M. Mechanisms of genotoxicity. A review of in vitro and in vivo studies with engineered nanoparticles. *Nanotoxicology* **8**, 233–278 (2014).

170. Nabiev, I., Mitchell, S., Davies, A., Williams, Y., Kelleher, D., Moore, R., Gun'ko Y. K. Nonfunctionalized nanocrystals can exploit a cell's active transport machinery delivering them to specific nuclear and cytoplasmic compartments. *Nano letters* **7**, 3452–3461 (2007).

171. Karlsson, H. L., Gliga, A. R., Calléja, F. M., Gonçalves, C. S., Wallinder, I. O., Vrieling, H., Fadeel, B. Mechanism-based genotoxicity screening of metal oxide nanoparticles using the ToxTracker panel of reporter cell lines. *Particle and Fibre Toxicology* **11**, 41 (2014).

172. Fu, P. P., Xia, Q., Hwang, H. M., Ray, P. C., Yu, H. Mechanisms of nanotoxicity: generation of reactive oxygen species. *Journal of Food and Drug Analysis* **22**, 64–75 (2014).

173. Lin, W., Xu, Y., Huang, C. C., Ma, Y., Shannon, K. B., Chen, D. R., Huang, Y. W. Toxicity of nano- and micro-sized ZnO particles in human lung epithelial cells. *Journal of Nanoparticle Research* **11**, 25–39 (2009).

174. Karlsson, H. L., Cronholm, P., Gustafsson, J., Moller, L. Copper oxide nanoparticles are highly toxic: a comparison between metal oxide nanoparticles and carbon nanotubes. *Chemical Research in Toxicology* **21**, 1726–1732 (2008).

175. Bhattacharya, K., Davoren, M., Boertz, J., Schins, R. P., Hoffmann, E., Dopp, E. Titanium dioxide nanoparticles induce oxidative stress and DNA-adduct formation but not DNA-breakage in human lung cells. *Particle and Fibre Toxicology* **6**, 17 (2009).

176. Falck, G. C. M., Lindberg, H. K., Suhonen, S., Vippola, M., Vanhala, E., Catalan, J., Savolainen K. Genotoxic effects of nanosized and fine TiO2. *Human and Experimental Toxicology* **28**, 339–352 (2009).

177. Rahman, Q., Lohani, M., Dopp, E., Pemsel, H., Jonas, L., Weiss, D. G., Schiffmann, D. Evidence that ultrafine titanium dioxide induces micronuclei and apoptosis in Syrian hamster embryo fibroblasts. *Environmental Health Perspectives* **110**, 797 (2002).

178. Huang, Y. W., Wu, C. H., Aronstam, R. S. Toxicity of transition metal oxide nanoparticles: recent insights from in vitro studies. *Materials* **3**, 4842–4859 (2010).

179. Oomen, A. G., Bos, P. M., Fernandes, T. F., Hund-Rinke, K., Boraschi, D., Byrne, H. J., Aschberger, K., et al. Concern-driven integrated approaches to nanomaterial testing and assessment–report of the NanoSafety Cluster Working Group 10. *Nanotoxicology* **8**, 334-348. (2014).

Appendices

Metal Oxide nanomaterials Genotoxicity Model Supplementary Information

Table S1 A. Assessment of the quality of the data points. Adherence to the minimum comet assay requirements (yellow boxes) and minimum physicochemical characterization (green boxes) of nanomaterials was evaluated by answering the questions in the header. Each data point represents a study reporting comet assay results for one or more metal oxides with the same or different core composition. The comet assay procedure and the characterization was done in the same way for nano metal oxides of the same data point. For a data point, genotoxicity results of the test may differ between the metal oxides with different core composition. If the question in the header was answered in the data point, then it is assigned by “Y” and if not by “N”.

Metal oxides core	Quality classification results (see Table 2)	The pH of unwinding: Alkaline, neutral, very alkaline.	Incubation with the enzymes: FPG, 8oxodG, Endo III.	Concentrations expressed in units?	Cytotoxicity tests performed?	Performed trend test for dose-response relationship?	Treatment time of at least 3h?	Performed comparison with either negative or positive?	Both negative and positive controls were used?	Demonstrated cellular uptake?	Any electron microscopy or atomic force microscopy image?	Chemical composition	Particle size / size distribution	Morphology/shape/form	Agglomeration/aggregation state	References	Genotoxic (+) and Non-genotoxic (−) assigned based on the WOE approach
CuO	1	Alkaline	None	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Wang et al., 2012	+

Fe_2O_3	1	Alkaline	None	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Auffan et al., 2006	-
TiO_2	1	Alkaline	None	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Barillet et al., 2010	+
TiO_2	1	Alkaline	FPG	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	N	Shukla et al., 2011	+
TiO_2	1	Alkaline	FPG	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Shukla et al., 2013	+
TiO_2	1	Neutral	None	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Saquib et al., 2012	+
TiO_2	1	Alkaline	EndoIII and 8oxoG	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Woodruff et al., 2012	+
ZnO	1	Alkaline	8oxodG	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	N	Valdiglesias et al., 2013	+
ZnO	1	Alkaline	None	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Sharma et al., 2009	+
MgO	1	-	-	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Mahmoud et al., 2016	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Roszak et al., 7	-
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Chen et al., Febbraio 5	-
TiO_2	2	Alkaline	FPG and EndoIII	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Reeves et al., 2008	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Rajapakse et al., 2013	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Prasad et al., 2013	+

ZnO	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Hackenberg et al., 2011b	+
Al_2O_3	2	Very alkaline	None	Y	Y	Y	Y	Y	Y	N	N	Y	Y	N	N	Kim et al., 2009	+
Bi_2O_3	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Liman, 2013	+
CeO_2	2	Very alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	De Marzi et al., 2013	+
SiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	N	Y	Y	Y	N	Choi et al., 2011	+
TiO_2	2	Neutral	FPG and EndoIII	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Petković et al., 2011a	+
ZnO	2	Alkaline	FPG and EndoIII	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Demir et al., 2014b	+
SiO_2	2	Alkaline	FPG	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Gehrke et al., 2011	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Bernardeschi et al., 2010	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Falck et al., 2009	+
TiO_2	2	Alkaline	FPG	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Petković et al., 2011b	+
TiO_2, ZnO	2	Alkaline	FPG	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Kermanizadeh et al., 2013	TiO_2 (+), ZnO (+)
ZnO	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Alarifi et al., 2013c	+

V_2O_3	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Lansiedel et al., 2009	+
V_2O_5	2	Alkaline	None	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Lansiedel et al., 2009	-
SiO_2	3	Alkaline	None	Y	N	Y	Y	Y	Y	N	N	Y	Y	Y	N	Wang et al., 2007a	-
TiO_2, ZnO	3	Alkaline	None	Y	N	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Demir et al., 2014a	TiO_2 (+), ZnO (+)
TiO_2, ZnO	2	Very alkaline	None	Y	N	Y	Y	Y	Y	N	N	Y	Y	N	N	Gopalan et al., 2009	TiO_2 (+), ZnO (+)
TiO_2	2	Alkaline	None	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	N	Prasad et al., 2014	+
Al_2O_3, TiO_2, ZrO_2	2	Alkaline	FPG and EndoIII	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Demir et al., 2013	Al_2O_3 (-), TiO_2 (+), ZrO_2 (-)
CuO	3	Alkaline	FPG and EndoIII	Y	Y	N	Y	Y	Y	N	Y	Y	Y	Y	Y	Di Bucchianico et al., 2013	+
ZnO	2	Alkaline	None	Y	Y	N	Y	Y	N	Y	Y	Y	Y	N	Y	Mu et al., 2014	+
TiO_2	2	Alkaline	8oxodG	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Jugan et al., 2012	+
CuO, TiO_2, ZnO	2	Alkaline	FPG	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Bayat et al., 2014	CuO (+), TiO_2 (+), ZnO (+)
SiO_2	2	Alkaline	None	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	N	Barnes et al., 2008	-
TiO_2	2	Alkaline	None	Y	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Hamzeh and Sunahara, 3	-

TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Kang et al., 2008	-
CuO, Fe_2O_3, Fe_3O_4, TiO_2	2	Alkaline	FPG	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Karlsson et al., 2009	CuO (+), Fe_2O_3 (-), Fe_3O_4 (-), TiO_2 (+)
SiO_2, ZnO	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Yang et al., 2009	SiO_2 (+), ZnO (+)
CuO, Fe_2O_3, Fe_3O_4, TiO_2, ZnO	2	Alkaline	FPG	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Karlsson et al., 2008	CuO (+), Fe_2O_3 (-), Fe_3O_4 (-), TiO_2 (+), CuO (-), ZnO (+)
Fe_3O_4	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	N	Ahamed et al., 2013	+
CuO	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Isani et al., Gennaio 11	-
Fe_2O_3	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Freyria et al., 2012	+
SiO_2	2	Very alkaline	None	Y	Y	Y	Y	Y	N	N	N	Y	Y	N	N	Wang et al., 2007b	-
TiO_2	2	Very alkaline	None	Y	Y	Y	Y	Y	N	N	N	Y	Y	N	N	Wang et al., 2007c	-
CeO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Kumari et al., 2014	+
CeO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Auffan et al., 2009	+
Co_3O_4	2	Very alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Alarifi et al., 2013b	+

Fe_3O_4	2	Alkaline	None	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	Y	Könczöl et al., 2011	+
SiO_2	2	Alkaline	8oxodG	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	Jin et al., 2007	-
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Ghosh et al., 2013	+
Fe_2O_3, Fe_3O_4, TiO_2	2	Alkaline	FPG	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Guichard et al., 2012	Fe_2O_3 (-), Fe_3O_4 (-), TiO_2 (+)
TiO2, ZnO	2	Alkaline	FPG	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Kermanizadeh et al., 2012	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	N	N	Ghosh et al., 2010	+
CuO	2	Alkaline	None	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Alarifi et al., 2013a	-
SiO_2	2	Alkaline	FPG and 8oxoG	Y	Y	Y	Y	Y	N	Y	N	Y	Y	N	Y	Gonzalez et al., 2010	-
TiO_2	2	Very alkaline	8oxodG	Y	Y	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Wan et al., 2012	-
SiO_2	2	Alkaline	FPG	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	Y	Y	Lankoff et al., 2013	-
ZnO	2	Alkaline	None	Y	Y	Y	N	Y	N	Y	Y	Y	Y	Y	N	Sarkar et al., 2014	+
TiO_2	2	Alkaline	None	Y	Y	Y	Y	Y	N	N	Y	Y	Y	Y	Y	Hackenberg et al., 2011a	-
MgO, SiO_2, TiO_2, ZnO	2	Alkaline	FPG	Y	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Y	Gerloff et al., 2009	MgO (-), SiO_2 (-), TiO_2 (+), ZnO (+)

Fe_3O_4	3	Alkaline	None	Y	N	N	Y	N	Y	Y	Y	Y	Y	Y	N	Gomaa et al., 2013	+
CeO_2	3	Alkaline	None	Y	N	Y	N	Y	Y	N	Y	Y	Y	Y	Y	Courbiere et al., 2013	+
TiO_2	3	Alkaline	FPG	Y	Y	N	N	Y	N	Y	Y	Y	Y	Y	N	Gurr et al., 2005	+
TiO_2	3	Alkaline	None	Y	Y	N	Y	Y	N	N	Y	Y	Y	N	N	Botelho et al., 2014	+
CuO	3	Alkaline	None	Y	Y	N	Y	Y	N	N	Y	Y	Y	Y	Y	Midander et al., 2009	+
Fe_2O_3, TiO_2	3	Alkaline	8oxodG	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	Y	Bhattacharya et al., 2009	Fe_2O_3 (-), TiO_2 (+)
SiO_2	3	Alkaline	None	Y	Y	Y	N	Y	N	N	Y	Y	Y	Y	Y	Gong et al., 2012	-
CeO_2	NR	Alkaline	None	Y	N	N	Y	Y	N	N	Y	Y	Y	N	N	Pierscionek et al., 2010	-
CeO_2, Co_3O_4, Fe_3O_4, NiO, SiO_2	NR	Alkaline	FPG	Y	N	N	Y	Y	N	N	Y	Y	Y	Y	Y	Kain et al., 2012	CeO_2(+), Co_3O_4 (+), Fe_3O_4 (+), NiO (+), SiO_2 (-)
SnO_2	NR	Alkaline	None	Y	N	Y	N	Y	N	N	Y	Y	Y	N	Y	Khan and Husain, 2014	-
Fe_3O_4	NR	–	–	–	–	–	–	–	–	–	–	–	–	–	–	Hong et al., 2011	excluded
TiO_2	NR	Alkaline	FPG and EndoIII	Y	N	N	Y	N	N	N	Y	Y	Y	Y	Y	Sekar et al., 2014	+

Table S1 B. Assessment of the size measurement and crystallographic characterization as was reported in the related literature sources: transmission electron microscopy (TEM), dynamic light scattering (DLS), SEM, PCS, expressed in nm, Brunette-Emmet- Teller (BET) and X ray diffraction analysis expressed in m2g-1, Metal oxides core Nominal size TEM analysis DLS analysis, SEM PCS BET specific surface area (m2g-1) X ray diffraction analysis (m2g-1) - Crystallographic shape is reported in the corresponding column.

References	Metal oxides core	Nominal size (nm)	Genotoxicity for each size reported in the same paper	TEM (nm)	DLS (nm)	SEM	PCS	BET (m^2g^{-1})	XRD (m^2g^{-1})	Crystallographic shape
Wang et al., 2012	CuO	10	+	20-40	276.4					
Auffan et al., 2006	Fe2O3	6	-	6						
Könczöl et al., 2011	Fe_3O_4	20-60	+				311			
Jin et al., 2007	SiO_2	50	-							
Wang et al., 2007a	SiO2	12.2	-							Quarz

Roszak et al., 2013	TiO_2	25	-		300			27.1		Mixture of rutile and anatase
Chen et al., 2014	TiO_2	75±15	-		473.6 and 486.5 (depending on the medium)					
Hackenberg et al., 2011a	TiO2	<25	-	285 ± 52						
Ghosh et al., 2013	TiO2	100	+	58.93±7.08	6180.73					
Botelho et al., 2014	TiO2	21 and <25	+		160.5 and 420.7					
Barillet et al., 2010	TiO_2	NR	+	12±3				17		Anatase (75%), Anatase (100%)

Gurr et al., 2005	TiO_2	10 and 20	+							Anatase
Kang et al., 2008	TiO_2	25	+	Between 15 and 30, agglomeration size 285±52				50		Mixture of Anatase (70%) and Rutile (85%)
Reeves et al., 2008	TiO2	5	+							Anatase
Shukla et al., 2011	TiO_2	NR	+	50	124.9 (water)					
Shukla et al., 2013	TiO_2	NR	+	30 to 70	124.9 (water) and 192.5 (medium)					Anatase
Rajapakse et al., 2013	TiO2	NR	+	15	820			190-290		Anatase
Saquib et al., 2012	TiO2	NR	+		13 (water), 152				30.6	Polyhedral rutile

					(mediu m)					
Prasad et al., 2013	TiO_2	27.5	+					49		Mixture of anatase (86%) and rutile (14%)
Hackenberg et al., 2011b	ZnO	<100	+	86 ± 41 x 42 ± 21						
Karlsson et al., 2009	CuO	42	+	Between 20 and 40	200					
Karlsson et al., 2009	Fe_2O_3	29	-	30-60	1600					
Karlsson et al., 2009	Fe_3O_4	29	-	30-60	1600					
Karlsson et al., 2009	TiO_2	63	+	20 to 100	300					
Yang et al., 2009	SiO2	NR	+	20.2±6.4						

Yang et al., 2009	ZnO	NR	+	20.2 ± 6.4						
Demir et al., 2013	Al2O3	16.7	-		16.7±1.3					
Demir et al., 2013	TiO2	2.3	+		1.8-2.8					
Demir et al., 2013	ZrO_2	6	-	6 ± 0.8						
Pierscionek et al., 2010	CeO2	NR	-	5.5					6.3	
Kain et al., 2012	CeO_2	<25	+	4-25	225					
Kain et al., 2012	Fe_3O_4	30	+	20-40	200					
Kain et al., 2012	NiO	<50	+	Between 2 and 67	167					

Kain et al., 2012	SiO_2	15	-	Between 11 and 27	8.7					
Midander et al., 2009	CuO	28	+			50				
Karlsson et al., 2008	CuO	42	+	20-40	220					
Karlsson et al., 2008	Fe_2O_3	29	-	30 and 60	1580					
Karlsson et al., 2008	Fe_3O_4	20-30	-	20-40	200					
Karlsson et al., 2008	ZnO	71	+	20-200	320					
Di Bucchianico et al., 2013	CuO	NR	in RAW 264.7 + in PBL cells -	7±1 (spheres), 7±1 x 40±10 (rodes), 1200±250 x 270±50 x						

				30±10 (spindles)						
Bayat et al., 2014	CuO	<50	+		1511±468 (water), 3475±357 (medium)					Monoclinic crystals
Bayat et al., 2014	TiO2	3-17	+		99.20±6.2					Rutile
Bayat et al., 2014	ZnO	<100	+		612±10.9 (water) 5294 ± 3184 (medium)					Hexagonal wurtzite
Guichard et al., 2012	Fe2O3	NR	-	35±14	900 (water)			39		
Guichard et al., 2012	Fe3O4	NR	-	27±8	Between 700			40		

					and 800					
Guichard et al., 2012	TiO2	NR	+	14 ± 4 and 25 ± 6				149		
Bhattacharya et al., 2009	Fe_2O_3	90	-	93	68					
Ahamed et al., 2013	Fe3O4	NR	+	24.83	247 (water), 213 (medium)					
Gomaa et al., 2013	Fe3O4	NR	+	8±2						
Gerloff et al., 2009	MgO	8	-					200		
Gerloff et al., 2009	SiO_2*	50±3	-							

Gerloff et al., 2009	TiO_2	20-80	+					50		Mixture of anatase (80%)
Gerloff et al., 2009	ZnO	10 and 20	+					70 and 50		
Barnes et al., 2008	SiO_2	Five different samples of nominal size : 30-400	-							Amorphous
Kermanizadeh et al., 2012	TiO2	Different samples of 7, 10 and 94	+	4-8/50-100, 80-400, 1-4/100/100-200	Different samples: 185, 742, 203-1487, 339					
Kermanizadeh et al., 2012	ZnO	100	+	20-250/50-350	306					
Demir et al., 2014a	TiO2	21 and 50	+		21 ± 0.8 and 50 ± 0.5					Respectively the two samples are anatase and mixture of anatase and rutile

Demir et al., 2014a	ZnO	≤35 and 50	+		35±1.1					
Sekar et al. (2014)	TiO2	NR	+			10–20 (anatase), 20-150 (anatase and rutile)		132.73 (anatase), 20.75 (anatase and rutile)		Anatase and rutile
Mu et al., 2014	ZnO	NR	+	40±20					90 - 160	Uncoated zincite
Ghosh et al., 2010	TiO_2	100	+							
Courbiere et al., 2013	CeO2	3	+		350					
Kim et al., 2009	Al_2O_3	<50	+							
Liman, 2013	Bi2O3	90-210	+							

De Marzi et al., 2013	CeO_2	40	+			16 to 22 nm.				
Isani et al., Gennaio 11	CuO	NR	+		500 ± 20 nm	20 to 200 nm				
Alarifi et al., 2013a	CuO	10	+	Between 20 and 40	276.4 (water)					
Freyria et al., 2012	Fe2O3	<100	+		50 (water)					
Wang et al., 2007b	SiO2	NR	-							Crystalline
Choi et al., 2011	SiO2	10	+							
Gonzalez et al., 2010	SiO_2	25	-	16.4± 2.5 and 60.4± 8.3	110 and 70					Crystalline
Gopalan et al., 2009	TiO2	NR	+			40-70				Anatase

Gopalan et al., 2009	ZnO	NR	+			40-70				
Wang et al., 2007c	TiO2	NR	+		6.57 nm; 8.2 nm; 196.52 nm					Crystalline
Woodruff et al., 2012	TiO2	NR	-	10x30						Anatase
Wan et al., 2012	TiO_2	28	-		280					Anatase (90%) and rutile (10%)
Petković et al., 2011a	TiO_2	<25 and <100	+					129.3 and 116.7		Rutile and anatase
Valdiglesias et al., 2013	ZnO	100	+		243.7			100		
Sarkar et al., 2014	ZnO	NR	-	45-150 (75 ± 5 average diameter)	45-150					

Demir et al., 2014b	ZnO	35 nm; 50-80 nm	+		36.42; 50.75					
Kumari et al., 2014	CeO_2	<25	+	25±1.512	269.7					
Auffan et al., 2009	CeO_2	7	+		15					
Alarifi et al., 2013b	Co3O4	NR	+	21	264.8					?
Lankoff et al., 2013	SiO2	NR	-			Average size 10 to 50		640-260		SiO2 with three types of functionlaisation, amorphous
Gong et al., 2012	SiO2	15, 30, 100	+	14.6; 20.4; 169.2	14.6±0.3, 20.4±1.7 and 169.2±3.1 (Medium)					Amorphous
Gehrke et al., 2011	SiO2	12; 40	-	16-40 and 50-100	165 and 271			200; 50; 4		Amorphous

Khan and Husain, 2014	SnO_2	NR	-	25-32					11.2	Crystalline
Bernardeschi et al., 2010	TiO2	<25	+							Anatase
Falck et al., 2009	TiO2	<25	+					222		Anatase (99.7%)
Jugan et al., 2012	TiO2	12, 20, 25	+	12, 21, 24				92; 73; 46		Anatase and rutile
Hamzeh and Sunahara, 2013	TiO2	5.9, 34.1, 15.5, 1-10	+		460, 400, 420, 600					Anatase and rutile
Prasad et al., 2014	TiO2	27.5	+					49		Anatase (86%) and rutile (14%)
Petković et al., 2011b	TiO2	<25	+					129.3	18	Anatase
Kermanizadeh et al., 2013	TiO2	NR	+	4-8/50-100, 80-400, 80-						

				400, 80-400						
Sharma et al., 2009	ZnO	NR	+	17	263					
Alarifi et al., 2013c	ZnO	NR	+	20-250/50-350				14	70 to > 100	Zincite
Kermanizadeh et al., 2013	ZnO	NR	+	20–200/10–450 and 20–200/ 10–450				14 and 18	70 to >100 and 58–93 nm	
Kain et al., 2012	Co_3O_4	<50	+	9-62	222					
Landsiedel, 2009	V_2O_3	<50	+					Average diameter 25 nm Length 100 – 1.000 nm		
Landsiedel, 2009	V_2O_5	70	-					Rod-shaped, spherical		

								diam.170 –180 nm		

*luminescent silica nanoparticles

Table S2. Calculated values of the MOPAC descriptors reported in Table 6 (page 70), for all the nanomaterials reported in Table 5 (page 69)

Metal oxides	HF (KCAL/MOL)	TE (eV)	EE (eV)	CORE (eV)	COSMO-SA (A°2)	COSMO (A°3)	IP (eV)	HOMO (eV)	LUMO (eV)	N.Fl (adimensional)	MW (g/mol)
Al2O3	-5162.20	-15398.50	-175859.00	160460.10	679.16	1280.89	7.32	-7.32	-1.37	180	1529.42
Bi2O3	-277.49	-17963.90	-234375.00	216410.80	852.64	1858.52	6.43	-6.44	-2.48	224	7455.34
CeO2	-4031.46	-20269.60	-286812.00	266542.40	878.68	1998.51	7.07	-7.07	-0.29	240	5507.67
Co3O4	-445.89	-19078.70	-265366.00	246287.20	517.86	883.12	7.36	-7.36	-1.90	204	1926.38
CuO	-1947.87	-46652.20	-967883.00	921230.80	619.14	1027.11	6.84	-6.84	-3.62	408	3818.18
Fe2O3	-5235.58	-57293.30	-1510297.00	1453004.00	1003.65	2418	8.51	-8.51	-5.22	578	5429.54
Fe3O4	-1300.94	-19023.70	-233472.00	214448.10	538.81	968.38	8.76	-8.76	-4.39	192	1852.31
MgO	-3334.27	-10296.80	-101052.00	90755.60	516.44	948.36	7.29	-7.29	-3.84	128	1289.74
NiO	-1783.99	-21622.30	-415687.00	394064.50	462.73	749.07	3.46	-3.46	0.83	256	2390.7
SiO2	-10558.30	-44533.20	-1040743.00	996210.20	1306.9	2961.68	7.74	-7.74	-3.67	512	3845.4
SnO2	-3521.02	-32654.60	-652905.00	620250.90	1043	2362.53	6.94	-6.94	-1.54	384	7233.06
TiO2	-4946.57	-21549.70	-320025.00	298475.60	823.4	1688.85	6.57	-6.57	-3.61	256	2556.76
V2O3	-4329.18	-27328.50	-540060.00	512731.70	862.78	1869.59	4.94	-4.94	-1.80	336	3597.15
V2O5	-2379.90	-13846.20	-158034.00	144188.10	501.4	816.4	6.05	-6.05	-1.90	160	1455.04
ZnO	-1666.65	-11658.30	-108506.00	96847.25	688.93	1058.82	6.10	-6.10	-2.52	144	2929.66
ZrO2	-6591.16	-20819.40	-325388.00	304568.80	824.22	1875.04	2.18	-2.18	1.52	250	3669.34

Table S3. Scaled values of the quantum-chemically calculated descriptors of the metal oxide nanomaterials data set and the scale attributes

Metal oxides core	HF (kcal/mol)	TE (eV)	EE (eV)	CORE (eV)	COSMO -SA (A°2)	COSMO (A°3)	IP (eV)	HOMO (eV)	LUMO (eV)	N.Fl (adimensional)	MW (g/mol)
Al_2O_3	-1.1367	-0.5281	-0.2843	0.2719	0.8312	0.7381	1.0607	-1.0606	-0.4633	0.5696	0.3691
Bi_2O_3	-0.0611	-0.6161	-0.3789	0.3666	1.0435	1.0710	0.9318	-0.9331	-0.8387	0.7088	1.7991
CeO_2	-0.8877	-0.6952	-0.4636	0.4516	1.0754	1.1516	1.0245	-1.0244	-0.0981	0.7595	1.3291
Co_3O_4	-0.0982	-0.6544	-0.4290	0.4173	0.6338	0.5089	1.0665	-1.0664	-0.6425	0.6455	0.4649
CuO	-0.4289	-1.6001	-1.5646	1.5608	0.7578	0.5919	0.9912	-0.9911	-1.2242	1.2911	0.9214
Fe_2O_3	-1.1529	-1.9651	-2.4415	2.4617	1.2284	1.3934	1.2332	-1.2331	-1.7653	1.8290	1.3102
Fe_3O_4	-0.2865	-0.6525	-0.3774	0.3633	0.6594	0.5580	1.2694	-1.2693	-1.4846	0.6076	0.4470
MgO	-0.7342	-0.3532	-0.1634	0.1538	0.6321	0.5465	1.0564	-1.0563	-1.2986	0.4050	0.3112
NiO	-0.3928	-0.7416	-0.6720	0.6676	0.5663	0.4316	0.5014	-0.5013	0.2807	0.8101	0.5769
SiO_2	-2.3250	-1.5274	-1.6824	1.6878	1.5995	1.7067	1.1216	-1.1215	-1.2411	1.6202	0.9280
SnO_2	-0.7753	-1.1200	-1.0555	1.0508	1.2765	1.3614	1.0057	-1.0056	-0.5208	1.2151	1.7455
TiO_2	-1.0893	-0.7391	-0.5173	0.5057	1.0078	0.9732	0.9520	-0.9520	-1.2208	0.8101	0.6170
V_2O_3	-0.9533	-0.9373	-0.8730	0.8687	1.0560	1.0773	0.7158	-0.7158	-0.6087	1.0632	0.8681
V_2O_5	-0.5241	-0.4749	-0.2555	0.2443	0.6137	0.4704	0.8767	-0.8766	-0.6425	0.5063	0.3511
ZnO	-0.3670	-0.3999	-0.1754	0.1641	0.8432	0.6101	0.8839	-0.8839	-0.8522	0.4557	0.7070
ZrO_2	-1.4514	-0.7141	-0.5260	0.5160	1.0088	1.0805	0.3159	-0.3159	0.5140	0.7911	0.8855
Scale Attribute	4541.26	29155.80	618599.8	590239.50	817.06	1735.37	6.90	6.90	2.96	316.02	4143.91

Table S4. The number of total studies evaluated, studies with size range 5-100 nm, categorized in each reliability class and the overall genotoxicity for each metal oxide nanoparticles.

Metal oxides core	All studies	studies with sizes 5-100 Nm	Class 1	Class 2	Class 3	Class 4	Positive results	Negative results	Overall genotox-icity
Al2O3	2	2		2			1	1	+
NiO	1	1				1	1		+
Co3O4	2	2		1		1	2		+
CuO	8	6		5	1		5	1	+
Fe2O3	6	6	1	4	1		1	5	-
Fe3O4	7	7		5	1	1	4	3	+
TiO_2	35	12	3	8	1		8	4	+
ZnO	15	15	2	13			15		+
SiO_2	12	12		10	1	1	3	9	-
V_2O_3	1	1		1			1		+
V_2O_5	1	1		1				1	-
MgO	1	1		1				1	-
ZrO_2	1	1		1				1	-
CeO_2	6	5		3		2	4	1	+
Bi_2O_3	1	1		1			1		+
SnO_2	1	1				1		1	-

Reference

Ahamed, M., A Alhadlaq, H., Alam, J., Khan, M., Ali, D., and Alarafi, S. (2013). Iron oxide nanoparticle-induced oxidative stress and genotoxicity in human skin epithelial and lung epithelial cell lines. Curr. Pharm. Des. 19, 6681–6690.

Alarifi, S., Ali, D., Verma, A., Alakhtani, S., and Ali, B.A. (2013a). Cytotoxicity and genotoxicity of copper oxide nanoparticles in human skin keratinocytes cells. Int. J. Toxicol. 1091581813487563.

Alarifi, S., Ali, D., Al Omar Suliman, Y., Ahamed, M., Siddiqui, M.A., and Al-Khedhairy, A.A. (2013b). Oxidative stress contributes to cobalt oxide nanoparticles-induced cytotoxicity and DNA damage in human hepatocarcinoma cells. Int. J. Nanomedicine 8, 189.

Alarifi, S., Ali, D., Alkahtani, S., Verma, A., Ahamed, M., Ahmed, M., and Alhadlaq, H.A. (2013c). Induction of oxidative stress, DNA damage, and apoptosis in a malignant human skin melanoma cell line after exposure to zinc oxide nanoparticles. Int. J. Nanomedicine 8, 983.

Auffan, M., Decome, L., Rose, J., Orsiere, T., De Meo, M., Briois, V., Chaneac, C., Olivi, L., Berge-Lefranc, J.L., Botta, A., et al. (2006). In vitro interactions between DMSA-coated maghemite nanoparticles and human fibroblasts: A physicochemical and cyto-genotoxical study. Environ. Sci. Technol. 40, 4367–4373.

Auffan, M., Rose, J., Orsiere, T., De Meo, M., Thill, A., Zeyons, O., Proux, O., Masion, A., Chaurand, P., Spalla, O., et al. (2009). CeO2 nanoparticles induce DNA damage towards human dermal fibroblasts in vitro. Nanotoxicology 3, 161–171.

Barillet, S., Simon-Deckers, A., Herlin-Boime, N., Mayne-L'Hermite, M., Reynaud, C., Cassio, D., Gouget, B., and Carrière, M. (2010). Toxicological consequences of TiO2, SiC nanoparticles and multi-walled carbon nanotubes exposure in several mammalian cell types: an in vitro study. J. Nanoparticle Res. 12, 61–73.

Barnes, C.A., Elsaesser, A., Arkusz, J., Smok, A., Palus, J., Leśniak, A., Salvati, A., Hanrahan, J.P., De Jong, W.H., Dziubaltowska, E., et al. (2008). Reproducible cornet assay of amorphous silica nanoparticles detects no genotoxicity. Nano Lett. 8, 3069–3074.

Bayat, N., Rajapakse, K., Marinsek-Logar, R., Drobne, D., and Cristobal, S. (2014). The effects of engineered nanoparticles on the cellular structure and growth of Saccharomyces cerevisiae. Nanotoxicology 8, 363–373.

Bernardeschi, M., Guidi, P., Scarcelli, V., Frenzilli, G., and Nigro, M. (2010). Genotoxic potential of TiO2 on bottlenose dolphin leukocytes. Anal. Bioanal. Chem. 396, 619–623.

Bhattacharya, K., Davoren, M., Boertz, J., Schins, R.P.F., Hoffmann, E., and Dopp, E. (2009). Titanium dioxide nanoparticles induce oxidative stress and DNA-adduct formation but not DNA-breakage in human lung cells. Part. Fibre Toxicol. 6.

Botelho, M.C., Costa, C., Silva, S., Costa, S., Dhawan, A., Oliveira, P.A., and Teixeira, J.P. (2014). Effects of titanium dioxide nanoparticles in human gastric epithelial cells in vitro. Biomed. Pharmacother. 68, 59–64.

Chen, Z., Wang, Y., Ba, T., Li, Y., Pu, J., Chen, T., Song, Y., Gu, Y., Qian, Q., Yang, J., et al. (Febbraio 5). Genotoxic evaluation of titanium dioxide nanoparticles in vivo and in vitro. Toxicol. Lett. 226, 314–319.

Choi, H.S., Kim, Y.J., Song, M., Song, M.K., and Ryu, J.C. (2011). Genotoxicity of nano-silica in mammalian cell lines. Toxicol. Environ. Health Sci. 3, 7–13.

Courbiere, B., Auffan, M., Rollais, R., Tassistro, V., Bonnefoy, A., Botta, A., Rose, J., Orsière, T., and Perrin, J. (2013). Ultrastructural Interactions and Genotoxicity Assay of Cerium Dioxide Nanoparticles on Mouse Oocytes. Int. J. Mol. Sci. 14, 21613–21628.

De Marzi, L., Monaco, A., De Lapuente, J., Ramos, D., Borras, M., Di Gioacchino, M., Santucci, S., and Poma, A. (2013). Cytotoxicity and genotoxicity of ceria nanoparticles on different cell lines in vitro. Int. J. Mol. Sci. 14, 3065–3077.

Demir, E., Burgucu, D., Turna, F., Aksakal, S., and Kaya, B. (2013). Determination of TiO2, ZrO2, and Al2O3 Nanoparticles on Genotoxic Responses in Human Peripheral Blood Lymphocytes and Cultured Embyronic Kidney Cells. J Toxicol Env. Health A 76, 990–1002.

Demir, E., Kaya, N., and KAYA, B. (2014a). Genotoxic effects of zinc oxide and titanium dioxide nanoparticles on root meristem cells of< i> Allium cepa by comet assay. Turk. J. Biol. 38, 31–39.

Demir, E., Akça, H., Kaya, B., Burgucu, D., Tokgün, O., Turna, F., Aksakal, S., Vales, G., Creus, A., and Marcos, R. (2014b). Zinc oxide nanoparticles: Genotoxicity, interactions with UV-light and cell-transforming potential. J. Hazard. Mater. 264, 420–429.

Di Bucchianico, S., Fabbrizi, M.R., Misra, S.K., Valsami-Jones, E., Berhanu, D., Reip, P., Bergamaschi, E., and Migliore, L. (2013). Multiple cytotoxic and genotoxic effects induced in vitro by differently shaped copper oxide nanomaterials. Mutagenesis 28, 287–299.

Falck, G.C., Lindberg, H.K., Suhonen, S., Vippola, M., Vanhala, E., Catalan, J., Savolainen, K., and Norppa, H. (2009). Genotoxic effects of nanosized and fine TiO2. Hum Exp Toxicol 28, 339–352.

Fischer K.B., Genotoxicity of Synthetic Nanomaterials, dissertation, 2008.

Freyria, F.S., Bonelli, B., Tomatis, M., Ghiazza, M., Gazzano, E., Ghigo, D., Garrone, E., and Fubini, B. (2012). Hematite nanoparticles larger than 90 nm show no sign of toxicity in terms of lactate dehydrogenase release, nitric oxide generation, apoptosis, and comet assay in murine alveolar macrophages and human lung epithelial cells. Chem. Res. Toxicol. 25, 850–861.

Gehrke, H., Frühmesser, A., Pelka, J., Esselen, M., Hecht, L.L., Blank, H., Schuchmann, H.P., Gerthsen, D., Marquardt, C., and Diabaté, S. (2011). In vitro toxicity of amorphous silica nanoparticles in human colon carcinoma cells. Nanotoxicology 7, 274–293.

Gerloff, K., Albrecht, C., Boots, A.W., Frster, I., and Schins, R.P.F. (2009). Cytotoxicity and oxidative DNA damage by nanoparticles in human intestinal Caco-2 cells. Nanotoxicology 3, 355–364.

Ghosh, M., Bandyopadhyay, M., and Mukherjee, A. (2010). Genotoxicity of titanium dioxide (TiO< sub> 2) nanoparticles at two trophic levels: Plant and human lymphocytes. Chemosphere 81, 1253–1262.

Ghosh, M., Chakraborty, A., and Mukherjee, A. (2013). Cytotoxic, genotoxic and the hemolytic effect of titanium dioxide (TiO2) nanoparticles on human erythrocyte and lymphocyte cells in vitro. J. Appl. Toxicol. 33, 1097–1110.

Gomaa, I.O., Kader, M.H.A., Eldin, T.A.S., and Heikal, O.A. (2013). Evaluation of in vitro mutagenicity and genotoxicity of magnetite nanoparticles. Drug Discov. Ther. 7, 116–123.

Gong, C., Tao, G., Yang, L., Liu, J., He, H., and Zhuang, Z. (2012). The role of reactive oxygen species in silicon dioxide nanoparticle-induced cytotoxicity and DNA damage in HaCaT cells. Mol. Biol. Rep. 39, 4915–4925.

Gonzalez, L., Thomassen, L.C., Plas, G., Rabolli, V., Napierska, D., Decordier, I., Roelants, M., Hoet, P.H., Kirschhock, C.E., and Martens, J.A. (2010). Exploring the aneugenic and clastogenic potential in the nanosize range: A549 human lung carcinoma cells and amorphous monodisperse silica nanoparticles as models. Nanotoxicology 4, 382–395.

Gopalan, R.C., Osman, I.F., Amani, A., De Matas, M., and Anderson, D. (2009). The effect of zinc oxide and titanium dioxide nanoparticles in the Comet assay with UVA photoactivation of human sperm and lymphocytes. Nanotoxicology 3, 33–39.

Guichard, Y., Schmit, J., Darne, C., Gaté, L., Goutet, M., Rousset, D., Rastoix, O., Wrobel, R., Witschger, O., Martin, A., et al. (2012). Cytotoxicity and genotoxicity of nanosized and microsized titanium dioxide and iron oxide particles in syrian hamster embryo cells. Ann. Occup. Hyg. 56, 631–644.

Gurr, J.R., Wang, A.S.S., Chen, C.H., and Jan, K.Y. (2005). Ultrafine titanium dioxide particles in the absence of photoactivation can induce oxidative damage to human bronchial epithelial cells. Toxicology 213, 66–73.

Hackenberg, S., Friehs, G., Kessler, M., Froelich, K., Ginzkey, C., Koehler, C., Scherzed, A., Burghartz, M., and Kleinsasser, N. (2011a). Nanosized titanium dioxide particles do not induce DNA damage in human peripheral blood lymphocytes. Environ. Mol. Mutagen. 52, 264–268.

Hackenberg, S., Scherzed, A., Technau, A., Kessler, M., Froelich, K., Ginzkey, C., Koehler, C., Burghartz, M., Hagen, R., and Kleinsasser, N. (2011b). Cytotoxic, genotoxic and pro-inflammatory effects of zinc oxide nanoparticles in human nasal mucosa cells in vitro. Toxicol. In Vitro 25, 657–663.

Hamzeh, M., & Sunahara, G. I. (2013). In vitro cytotoxicity and genotoxicity studies of titanium dioxide (TiO2) nanoparticles in Chinese hamster lung fibroblast cells. Toxicology in Vitro, 27(2), 864-873.

Isani, G., Falcioni, M.L., Barucca, G., Sekar, D., Andreani, G., Carpenè, E., and Falcioni, G. (Gennaio 11). Comparative toxicity of CuO nanoparticles and CuSO4 in rainbow trout. Ecotoxicol. Environ. Saf. 97, 40–46.

Jin, Y., Kannan, S., Wu, M., and Zhao, J.X. (2007). Toxicity of luminescent silica nanoparticles to living cells. Chem. Res. Toxicol. 20, 1126–1133.

Jugan, M.-L., Barillet, S., Simon-Deckers, A., Herlin-Boime, N., Sauvaigo, S., Douki, T., and Carriere, M. (2012). Titanium dioxide nanoparticles exhibit genotoxicity and impair DNA repair activity in A549 cells. Nanotoxicology 6, 501–513.

Kain, J., Karlsson, H.L., and Möller, L. (2012). DNA damage induced by micro- and nanoparticles - interaction with FPG influences the detection of DNA oxidation in the comet assay. Mutagenesis 27, 491–500.

Kang, S.J., Kim, B.M., Lee, Y.J., and Chung, H.W. (2008). Titanium dioxide nanoparticles trigger p53-mediated damage response in peripheral blood lymphocytes. Env. Mol Mutagen 49, 399–405.

Karlsson, H.L., Cronholm, P., Gustafsson, J., and Möller, L. (2008). Copper oxide nanoparticles are highly toxic: A comparison between metal oxide nanoparticles and carbon nanotubes. Chem. Res. Toxicol. 21, 1726–1732.

Karlsson, H.L., Gustafsson, J., Cronholm, P., and Möller, L. (2009). Size-dependent toxicity of metal oxide particles-A comparison between nano- and micrometer size. Toxicol. Lett. 188, 112–118.

Kermanizadeh, A., Gaiser, B.K., Hutchison, G.R., and Stone, V. (2012). An in vitro liver model-assessing oxidative stress and genotoxicity following exposure of hepatocytes to a panel of engineered nanomaterials. Part Fibre Toxicol 9, 28.

Kermanizadeh, A., Vranic, S., Boland, S., Moreau, K., Baeza-Squiban, A., Gaiser, B.K., Andrzejczuk, L.A., and Stone, V. (2013). An in vitro assessment of panel of engineered nanomaterials using a human renal cell line: Cytotoxicity, pro-inflammatory response, oxidative stress and genotoxicity. BMC Nephrol. 14.

Khan, M.J., and Husain, Q. (2014). Influence of pH and temperature on the activity of SnO2-bound α-amylase: A genotoxicity assessment of SnO2 nanoparticles. Prep. Biochem. Biotechnol. 44, 558–571.

Kim, Y.-J., Choi, H.-S., Song, M.-K., Youk, D.-Y., Kim, J.-H., and Ryu, J.-C. (2009). Genotoxicity of aluminum oxide (Al2O3) nanoparticle in mammalian cell lines. Mol. Cell. Toxicol. 5, 172–178.

Könczöl, M., Ebeling, S., Goldenberg, E., Treude, F., Gminski, R., Gieré, R., Grobéty, B., Rothen-Rutishauser, B., Merfort, I., and Mersch-Sundermann, V. (2011). Cytotoxicity and genotoxicity of size-fractionated iron oxide (magnetite) in A549 human lung epithelial cells: role of ROS, JNK, and NF-κB. Chem Res Toxicol 24.

Kumari, M., Singh, S.P., Chinde, S., Rahman, M.F., Mahboob, M., and Grover, P. (2014). Toxicity Study of Cerium Oxide Nanoparticles in Human Neuroblastoma Cells. Int. J. Toxicol. 1091581814522305.

Lankoff, A., Arabski, M., Wegierek-Ciuk, A., Kruszewski, M., Lisowska, H., Banasik-Nowak, A., Rozga-Wijas, K., Wojewodzka, M., and Slomkowski, S. (2013). Effect of surface modification of silica nanoparticles on toxicity and cellular uptake by human peripheral blood lymphocytes in vitro. Nanotoxicology 7, 235–250.

Landsiedel, R., Kapp, M. D., Schulz, M., Wiench, K., & Oesch, F. (2009). Genotoxicity investigations on nanomaterials: methods, preparation and characterization of test material, potential artifacts and limitations—many questions, some answers. Mutation Research/Reviews in Mutation Research, 681(2), 241-258.

Liman, R. (2013). Genotoxic effects of Bismuth (III) oxide nanoparticles by Allium and Comet assay. Chemosphere 93, 269–273.

Mahmoud, A., Ezgi, Ö., Merve, A., & Özhan, G. (2016). In vitro toxicological assessment of magnesium oxide nanoparticle exposure in several mammalian cell types. International journal of toxicology, 35(4), 429-437.

Midander, K., Cronholm, P., Karlsson, H.L., Elihn, K., Möller, L., Leygraf, C., and Wallinder, I.O. (2009). Surface characteristics, copper release, and toxicity of nano- and micrometer-sized copper and copper(II) oxide particles: a cross-disciplinary study. Small 5, 389–399.

Mu, Q., David, C.A., Galceran, J., Rey-Castro, C., Krzemiński, L., Wallace, R., Bamiduro, F., Milne, S.J., Hondow, N.S., Brydson, R., et al. (2014). Systematic investigation of the physicochemical factors that contribute to the toxicity of ZnO nanoparticles. Chem Res Toxicol 27, 558–567.

Petković, J., Zegura, B., Stevanović, M., Drnovšek, N., Uskoković, D., Novak, S., and Filipič, M. (2011a). DNA damage and alterations in expression of DNA damage responsive genes induced by TiO2 nanoparticles in human hepatoma HepG2 cells. Nanotoxicology 5, 341–353.

Petković, J., Küzma, T., Rade, K., Novak, S., and Filipič, M. (2011b). Pre-irradiation of anatase TiO2 particles with UV enhances their cytotoxic and genotoxic potential in human hepatoma HepG2 cells. J Hazard Mater 196, 145–152.

Pierscionek, B.K., Li, Y., Yasseen, A.A., Colhoun, L.M., Schachar, R.A., and Chen, W. (2010). Nanoceria have no genotoxic effect on human lens epithelial cells. Nanotechnology 21.

Prasad, R.Y., Wallace, K., Daniel, K.M., Tennant, A.H., Zucker, R.M., Strickland, J., Dreher, K., Kligerman, A.D., Blackman, C.F., and Demarini, D.M. (2013). Effect of treatment media on the agglomeration of titanium dioxide nanoparticles: impact on genotoxicity, cellular interaction, and cell cycle. ACS Nano 7, 1929–1942.

Prasad, R.Y., Simmons, S.O., Killius, M.G., Zucker, R.M., Kligerman, A.D., Blackman, C.F., Fry, R.C., and Demarini, D.M. (2014). Cellular interactions and biological responses to titanium dioxide nanoparticles in HepG2 and BEAS-2B cells: role of cell culture media. Env. Mol Mutagen 55, 336–342.

Rajapakse, K., Drobne, D., Kastelec, D., and Marinsek-Logar, R. (2013). Experimental evidence of false-positive Comet test results due to TiO2 particle--assay interactions. Nanotoxicology 7, 1043–1051.

Reeves, J.F., Davies, S.J., Dodd, N.J., and Jha, A.N. (2008). Hydroxyl radicals (*OH) are associated with titanium dioxide (TiO(2)) nanoparticle-induced cytotoxicity and oxidative DNA damage in fish cells. Mutat Res 640, 113–122.

Roszak, J., Stępnik, M., Nocuń, M., Ferlińska, M., Smok-Pieniążek, A., Grobelny, J., Tomaszewska, E., Wąsowicz, W., and Cieślak, M. (7). A strategy for in vitro safety testing of nanotitania-modified textile products. J. Hazard. Mater. 256–257, 67–75.

Saquib, Q., Al-Khedhairy, A.A., Siddiqui, M.A., Abou-Tarboush, F.M., Azam, A., and Musarrat, J. (2012). Titanium dioxide nanoparticles induced cytotoxicity, oxidative stress and DNA damage in human amnion epithelial (WISH) cells. Toxicol Vitro 26, 351–361.

Sarkar, J., Ghosh, M., Mukherjee, A., Chattopadhyay, D., and Acharya, K. (2014). Biosynthesis and safety evaluation of ZnO nanoparticles. Bioprocess Biosyst Eng 37, 165–171.

Sekar, D., Falcioni, M.L., Barucca, G., and Falcioni, G. (2014). DNA damage and repair following In vitro exposure to two different forms of titanium dioxide nanoparticles on trout erythrocyte. Env. Toxicol 29, 117–127.

Sharma, V., Shukla, R.K., Saxena, N., Parmar, D., Das, M., and Dhawan, A. (2009). DNA damaging potential of zinc oxide nanoparticles in human epidermal cells. Toxicol. Lett. 185, 211–218.

Shukla, R.K., Sharma, V., Pandey, A.K., Singh, S., Sultana, S., and Dhawan, A. (2011). ROS-mediated genotoxicity induced by titanium dioxide nanoparticles in human epidermal cells. Toxicol Vitro 25, 231–241.

Shukla, R.K., Kumar, A., Gurbani, D., Pandey, A.K., Singh, S., and Dhawan, A. (2013). TiO(2) nanoparticles induce oxidative DNA damage and apoptosis in human liver cells. Nanotoxicology 7, 48–60.

Valdiglesias, V., Costa, C., Kiliç, G., Costa, S., Pásaro, E., Laffon, B., and Teixeira, J.P. (2013). Neuronal cytotoxicity and genotoxicity induced by zinc oxide nanoparticles. Environ. Int. 55, 92–100.

Wan, R., Mo, Y., Feng, L., Chien, S., Tollerud, D.J., and Zhang, Q. (2012). DNA damage caused by metal nanoparticles: involvement of oxidative stress and activation of ATM. Chem Res Toxicol 25, 1402–1411.

Wang, J.J., Wang, H., and Sanderson, B.J.S. (2007a). Ultrafine quartz-induced damage in human lymphoblastoid cells in vitro using three genetic damage end-points. Toxicol. Mech. Methods 17, 223–232.

Wang, J.J., Sanderson, B.J.S., and Wang, H. (2007b). Cytotoxicity and genotoxicity of ultrafine crystalline SiO2 participate in cultured human lymphoblastoid cells. Environ. Mol. Mutagen. 48, 151–157.

Wang, J.J., Sanderson, B.J.S., and Wang, H. (2007c). Cyto- and genotoxicity of ultrafine TiO2 particles in cultured human lymphoblastoid cells. Mutat. Res. - Genet. Toxicol. Environ. Mutagen. 628, 99–106.

Wang, Z., Li, N., Zhao, J., White, J.C., Qu, P., and Xing, B. (2012). CuO nanoparticle interaction with human epithelial cells: Cellular uptake, location, export, and genotoxicity. Chem. Res. Toxicol. 25, 1512–1521.

Woodruff, R.S., Li, Y., Yan, J., Bishop, M., Jones, M.Y., Watanabe, F., Biris, A.S., Rice, P., Zhou, T., and Chen, T. (2012). Genotoxicity evaluation of titanium dioxide nanoparticles using the Ames test and Comet assay. J. Appl. Toxicol. 32, 934–943.

Yang, H., Liu, C., Yang, D., Zhang, H., and Xi, Z. (2009). Comparative study of cytotoxicity, oxidative stress and genotoxicity induced by four typical nanomaterials: The role of particle size, shape and composition. J. Appl. Toxicol. 29, 69–78.

Journal of Environmental Science and Health, Part C

Environmental Carcinogenesis and Ecotoxicology Reviews

ISSN: 1059-0501 (Print) 1532-4095 (Online) Journal homepage: http://www.tandfonline.com/loi/lesc2

New clues on carcinogenicity-related substructures derived from mining two large datasets of chemical compounds

Azadi Golbamaki, Emilio Benfenati, Nazanin Golbamaki, Alberto Manganaro, Erinc Merdivan, Alessandra Roncaglioni & Giuseppina Gini

To cite this article: Azadi Golbamaki, Emilio Benfenati, Nazanin Golbamaki, Alberto Manganaro, Erinc Merdivan, Alessandra Roncaglioni & Giuseppina Gini (2016) New clues on carcinogenicityrelated substructures derived from mining two large datasets of chemical compounds, Journal of Environmental Science and Health, Part C, 34:2, 97-113, DOI: 10.1080/10590501.2016.1166879

To link to this article: https://doi.org/10.1080/10590501.2016.1166879

Accepted author version posted online: 17 Mar 2016.
Published online: 17 Mar 2016.

Submit your article to this journal

Article views: 127

View related articles

View Crossmark data

Citing articles: 3 View citing articles

JOURNAL OF ENVIRONMENTAL SCIENCE AND HEALTH, PART C
2016 VOL. 34,. NO. 2, 97-113
http://dx.doi.org/10.1080/10590501.2016.1166879

New clues on carcinogenicity-related substructures derived from mining two large datasets of chemical compounds

Azadi Golbamaki[a], Emilio Benfenati[a], Nazanin Golbamaki[b], Alberto Manganaro[a], Erinc Merdivan[c], Alessandra Roncaglioni[a], and Giuseppina Gini[d]

[a]Laboratory of Environmental Chemistry and Toxicology, IRCCS - Istituto di Ricerche Farmacologiche Mario Negri, Milan, Italy; [b]DRC/VIVA/METO Unit, Institut National de l.Environnement Industriel et des Risques (INERIS), Parc Technologique Alata, Verneuil en Halatte, France; [c]Faculty of Engineering and Natural Sciences, Sabancı University, Tuzla/Istanbul, Turkey; [d]DEIB, Politecnico di Milano, Milan, Italy

ABSTRACT
In this study, new molecular fragments associated with genotoxic and nongenotoxic carcinogens are introduced to estimate the carcinogenic potential of compounds. Two rule-based carcinogenesis models were developed with the aid of SARpy: model R (from rodents' experimental data) and model E (from human carcinogenicity data). Structural alert extraction method of SARpy uses a completely automated and unbiased manner with statistical significance. The carcinogenicity models developed in this study are collections of carcinogenic potential fragments that were extracted from two carcinogenicity databases: the ANTARES carcinogenicity dataset with information from bioassay on rats and the combination of ISSCAN and CGX datasets, which take into accounts human-based assessment. The performance of these two models was evaluated in terms of cross-validation and external validation using a 258 compound case study dataset. Combining R and H predictions and scoring a positive or negative result when both models are concordant on a prediction, increased accuracy to 72% and specificity to 79% on the external test set. The carcinogenic fragments present in the two models were compared and analyzed from the point of view of chemical class. The results of this study show that the developed rule sets will be a useful tool to identify some new structural alerts of carcinogenicity and provide effective information on the molecular structures of carcinogenic chemicals.

KEYWORDS
Carcinogenicity, QSAR, structural alerts, SARpy, *in silico*, molecular structures

Introduction

Identification, classification, and risk assessment of carcinogenic chemicals by international organizations and national agencies of health and safety have made remarkable progress in recent years. The European Commission (EC) substantially modified and replaced the Directive 67/548/EEC and 93/101/EEC with Regulation (EC) 1272/2008 on risks and hazards of carcinogens and mutagens [1]. The new regulation introduced the globally harmonized system of classification and labelling of chemicals. Under these directives experimental data studies on chemical carcinogens have been digitally collected with the aim of harmonizing national

measures on classification, packaging, and labelling of dangerous substances, to facilitate the establishment of a single market and to provide protection for public health and the environment. The new regulation complements the REACH regulation on the registration, evaluation, authorization, and restriction of chemicals.

Research has provided evidence that chemicals may cause cancer in animals and humans by one of several general mechanisms of action (MoA), generally classified into genotoxic and nongenotoxic. Genotoxic carcinogens cause damage to DNA, thus, many known mutagens are in this category, and often mutation is one of the first steps in the development of cancer [2]. Epigenetic or nongenotoxic carcinogens do not bind covalently to DNA, and are usually negative in the standard mutagenicity assays [3]. The unifying feature of all genotoxic carcinogens is that they are either electrophiles or can be activated to electrophilic reactive intermediates. On the contrary, nongenotoxic carcinogens act through a large variety of different and specific mechanisms.

For more than 35 years, many chemicals have been tested by government agencies, private companies, and research institutes using the two-year rodent carcinogenesis bioassay. Most of the chemicals or processes that have been associated with human carcinogenicity, as studied by epidemiological investigations, are shown to cause tumors in rats and mice [4–6]. However, all compounds shown to induce cancer in laboratory rats and mice are not necessarily human carcinogens [7].

In the past ten years, research into the MoA and carcinogenesis has increased and the relevance of the carcinogenicity findings in rodents to human risk has been investigated in many publications [8–10]. The results of research demonstrated that doses used in the bioassays may do not develop toxicity in humans exposed to same levels of these chemicals; in addition, rats and mice tumors occur in a sex, age, and strain- or stock-dependent manner. In consequence of these points, the regulatory agencies consider that the high occurrence of tumors in the standard twoyear rodent carcinogenesis bioassay is often not relevant to risk evaluation of human carcinogenesis [11]. Variability of the tumors in rodents is another problem of this assay. To deal with the problems of two-year rodent carcinogenesis bioassay alternative methods are suggested

by scientists and regulatory agencies. These methods include use of the toxicity level (LD50) in rodents, [12] *in vitro* cell transformation and other assays, *in silico* methods, or computerized prediction of carcinogenicity based on structure and chemical class [13] .Each method has its own strengths and weaknesses, and analysis of carcinogenicity of a specific chemical and its MoA in human is better to be assessed based on the weight of evidence.

Among the *in silico* methods, the use of various computational techniques such as (quantitative) structure-activity relationship ((Q)SAR) modelling is supported by several legislative authorities [14–16]. (Q)SAR models consist of mathematical relationships between physicochemical properties of chemicals and their biological activity, thus being able to calculate a quantitative value (for the activity) given the structure of a chemical. These mathematical relationships can be simple linear regression equations, or more complex nonlinear algorithms, and can be developed using several approaches such as neural networks, support vector machines, decision trees, and many others. Conversely, SAR identifies the differences of compounds in two categories (e.g., active or inactive) and predicts an untested compound as “toxic” in case it has a toxic potential or “nontoxic” if not. Overall, (Q)SAR models are useful for the prediction of toxicity of untested chemicals saving costs and the need for testing on animals [17,18].

Following the theory of electrophilic reactivity of (many) carcinogens of James and Elizabeth Millers, [19,20] the advancement of the knowledge of carcinogenic chemicals have received distinguished contributions from many scientists. The *salmonella typhimurium mutagenicity* assay by Bruce Ames [7] and the compilation of the lists of carcinogenic and mutagenic structural alerts (SA) by John Ashby [21] were two fundamental contributions to this field. SAs identified and collected by John Ashby’s are indeed reactive functional groups responsible for the induction of mutation or cancer, and are so-called genotoxic carcinogens. On the other hand, the *Salmonella* assay is the most predictive assay for genotoxic carcinogens and no other nongenotoxic mutagenicity test exists [22]. Despite the extensive knowledge of genotoxic SAs, the use of SAs for identifying nongenotoxic carcinogens is restricted. Nongenotoxic carcinogens use many different MoA and they lack an apparent unifying mechanism. According to this diversity, different (Q)SAR models have been developed and made available for analysis and

identification of SAs. A number of nongenotoxic SAs and their characteristics have been published in Woo and Lai [3].

One of the most recent rule sets defined by human expert for mutagenic carcinogenicity has been developed by Benigni and Bossa [23,24]. The updated version of this rule set [24] is implemented in Toxtree version 2.6.13,[25] a software application that investigates the presence of the genotoxic and nongenotoxic SAs in the chemical structures of the compounds. Alongside the rule-based (Q)SAR software that check the presence of human expert SAs in the chemical structures, there are statistically based (Q)SARs, which create models by using categorized active and inactive chemicals in a learning set to identify SAs that are associated with a particular toxicological activity. The high accuracy of the predictions performed by data mining and artificial intelligence has made these methods important tools to be used for preliminary research and for discovery of the mechanism of action that are still unknown. These methods, however, comparing to rule-based models are less transparent to the end user. Historically, the Computer Automated Structure Evaluation (CASE/MultiCASE) [26] program is a SAR expert system that identifies two-dimensional structural features or biophores, which can be used for the prediction of unknown compounds as potential toxins. This statistically based program does not use the knowledge on the mechanisms of action, but reanalyze the dataset of chemicals trying to link the structures of chemicals into their toxic activity. On the other hand, SAs developed by human experts were integrated in software such as OncoLogic [27] and DEREK [28].

In this study we used SARpy, [29] a commercially free statistically based program, for the extraction of potential carcinogenic SAs from two different learning sets. The approach that we have taken in developing the two new carcinogenicity models is mainly based on statistical evaluation of the chemicals in our learning sets categorized in two groups of carcinogens and noncarcinogens. The SARpy's method of identification of the SAs that are associated with a particular biological or toxicological activity does not demand a priori knowledge about MoA of the compounds and performs purely on a statistical basis. Two different carcinogenicity datasets have been prepared as learning sets and SARpy extracted two different models from these two datasets. The internal and external evaluation of the models have been assessed

thoroughly.Thechoiceoftakingintoconsiderationtwosubstantiallydifferentlearning sets and developing two models is due to different characterization of these data. The first dataset contains exclusively rodent carcinogenicity data based on presence of carcinogenic effects in male or female rats, while the second dataset takes into account human-based assessments and data retrieved from different assays. This suggests to obtain two different carcinogenicity models.

Finally, the SAs in the two rule sets are analyzed from the point of view of chemical class and the same SAs present in both rule sets are revised. The two developed models have been made available inside VEGA (http://www.vega-qsar.eu/), [30] an open source platform that already offers several (Q)SAR models.

Material and methods

Carcinogenesis data sources

ANTARES carcinogenicity dataset: Rat carcinogenesis learning set

Compounds for the first model's learning set were obtained from the carcinogenicity database of EU-funded project ANTARES [31]. The ANTAREs' carcinogenicity database is a collection of chemical rat carcinogenesis data (presence of carcinogenic effects in male or female rats) obtained from the EU-funded project CAESAR[32] dataset and the "FDA 2009 SAR Carcinogenicity—SAR Structures" database. The CAESAR toxicity values were originated from the Distributed Structure-Searchable Toxicity DSSTox database, which was built from the Lois Gold's Carcinogenic Potency Database [33]. The compounds with a definite TD50 (which is the dose that produces an increase of 50% of the tumors in animals) value for rat in this dataset were labeled as carcinogenic, while the remaining were labeled as noncarcinogenic. Additional 738 chemicals different from the 805 CAESAR compounds were added. The added chemicals are from the "FDA 2009 SAR Carcinogenicity— SAR Structures" database using the Leadscope database [34]. Here a categorical label for carcinogenicity was already contained in the original dataset and again the compound was labeled as carcinogenic if a positive outcome was detected in male or female rats. So a total number of 1543 compounds constituted the ANTARES dataset.

ISS carcinogenicity database and carcinogenicity genotoxicity experience dataset: Different species carcinogenesis learning set

The ISS Carcinogenicity (ISSCAN) database [35] is provided by the Istituto Superiore di Sanità (ISS). It is originally aimed at developing predictive models for carcinogenicity of chemicals. The great part of the chemicals in this database are classified as carcinogens by various regulatory agencies and scientific bodies. The database has been specifically designed as an expert decision support tool and contains information on chemicals tested with the long-term carcinogenicity bioassay on rodents (presence of carcinogenic effects in male or female rats and mice). This carcinogenicity dataset contains 622 carcinogens, 210 noncarcinogens and 58 equivocals.

Compounds for the second model's learning set were obtained by merging the ISSCAN database and the Carcinogenicity Genotoxicity eXperience (CGX) database. More information on the CGX database can be found in Kirkland and colleagues [36]. In this study, compounds used for development of the new models had to be either positive or negative; thus, compounds with equivocal results in the databases have been removed. In particular, from the original ISSCAN dataset with 890 compounds, we removed 58 compounds, while the CGX database did not contain any equivocal result.

All compounds in the combined dataset have been checked for their consistency between the two sources. We found 651 compounds in common, 15 of them with inconsistent carcinogenicity values. These compounds have been removed from the combined dataset.

Comparison with the ANTARES dataset

We compared the final list of compounds with the ANTARES carcinogenicity dataset prepared for the development of the first model. We found 105 compounds with conflicting values when compared with the compounds in the ANTARES dataset. In order to develop a more conservative model, we opted to remove only 15 compounds which had positive result in the ANTARES dataset and negative results in the combined second dataset, and left as carcinogenic those that had carcinogenicity result the opposite way. Consequently, there are 90 positive compounds in the combined database which are negative in the ANTARES dataset. Afterward,

we checked and cleaned the structures manually, and by the help of the istMolBase [37] and InstantJChem [38] software formed the final dataset.

In addition, the compounds have been checked for their molecular structure. We adopted only the substances with connected molecular structure; those which had unconnected structures have been removed from the dataset. The overall dataset consisted of 986 compounds with 734 carcinogens and 252 noncarcinogens. Each compound in the list had a chemical name, a CAS number, a Simplified Molecular Input Line Entry Specification (SMILES), [39] and its categorical designation (i.e., carcinogen or noncarcinogen). In the present study, this combined dataset is conventionally called ISSCAN-CGX.

Data for model validation

ECHA database

We prepared an external test set for the validation of the developed models from carcinogenicity the eChemPortal inventory [40]. For this purpose, we made two queries on this database. The first query contained the following restrictions:

- Study result type: experimental result
- Reliability: 1 and 2
- Species: mouse and rat • Maximum number of studies: 4 The second query consisted of:
- Study result type: experimental result
- Reliability: 1 and 2
- Species: mouse and rat
- Sources: any guideline and exposure route

The list resulted from the first query comprised 308 compounds, whereas the second query returned a list of 166 compounds, which were mostly in common with the results of the first query. The studies conducted for the first list of compounds have been manually evaluated. Afterward, we looked into the Classification Labelling and Packaging (CLP) inventory [41] for the positive (i.e., carcinogenic) chemicals collected by the previous queries. Inside the CLP inventory we found 68 compounds, which were already present in our data collection. The latter search confirmed the carcinogenic property of these compounds.

The dataset consisted of 64 positive compounds, 169 negative compounds, and 90 equivocal compounds. The equivocal results are due to the presence of conflicting information in different sources or different studies in the same source.

It should be noticed that for already classified compounds (no conflicting information), the level of uncertainty in the assignment is not homogeneous, because some of the compounds were classified on the basis of a single study (i.e., data present in one single source).

From the reliability point of view, in the data collected in our dataset, 49 positive compounds have positive carcinogenic effect in at least two sources. Fifty-seven negative compounds are noncarcinogenic in both lists, and they are not present in the list of compounds retrieved from the CLP inventory. Sixty-four compounds are considered as noncarcinogens because of the presence of only one single study in the two lists.

SARpy

The SAR in Python (SARpy) program is a Python script based on the OpenBabel chemical library. SARpy creates classification models by using categorized active and inactive chemicals in a learning set to identify molecular fragments that are associated with a particular biological, pharmaceutical, or toxicological activity. The algorithm generates molecular substructures of arbitrary complexity, and the fragments candidates to become SAs are automatically selected on the basis of their prediction performance in a learning set.

The output of SARpy consists in a set of rules in the form:

IF contains THEN,

where the SA is expressed as a SMARTS string, for use by human experts or other chemical software. SMARTS notations are text representations of substructures [36] that allow specification of wildcard atoms and bonds, which can be used to formulate substructure queries for a chemical database. Those rules can be used as a predictive model simply by calling a SMARTS matching program. For the matching phase, SMILES and the SMARTS strings are translated into graphs and the two graphs are compared to each other [42].

Extracting active fragments

R (rat) model

To obtain a more comprehensive collection of potential carcinogenic fragments, five learning sets were randomly created from the ANTARES carcinogenicity dataset with 1543 compounds, preserving 80% for the learning set and 20% for the evaluation set. In other words, for each model a random set of 20% of chemicals in the learning set was removed, with the remaining 80% of the compounds a model was developed and the activity of the compounds left out was predicted with the same model. We combined the five models and put together the lists of the potential active fragments, removed the duplicates and eliminated the SAs with likelihood ratio lower than two. We opted for the likelihood ratio threshold of two in order to retain the SAs that are statistically more significant. A measure of each fragment's association with biological activity is determined by SARpy as "training likelihood ratio," and it is given along with the list of the potential fragments or the rule set in the output. The likelihood ratio can be taken into account to determine the goodness of a SA identified by SARpy. Even if a SA that is associated with activity (i.e., carcinogenicity) is present in a molecular structure, the molecule may contain other fragments that make it inactive (i.e., noncarcinogen), thus the specific SA might not be expected to be found only in active compounds. This evidence is the basis of the determination of the likelihood ratio.

Using the SARpy software, each chemical in the learning set was fragmented *in silico* into all possible fragments meeting user-specified criteria. For this study we extracted only the "ACTIVE" fragments (or SAs) and the default values for the minimum and maximum number of atoms in a fragment were set for the fragment extractions of each model (minimum = 2; maximum = 18). Another configuration to establish by the user is the minimum number of compounds in the learning set in which an active (or inactive) fragment is found. In our analysis, the minimum number of compounds that contain a potential active fragment was set to three. Conventionally, in this study we call this model R.

E (expert) model

SARpy was used for model development and statistical analysis using the ISSCANCGX dataset.

The extraction settings are as follows: the minimum number of atoms in a fragment is equal to four, whereas the maximum number of atoms is equal to 10, and the minimum number of compounds containing the active fragment is six. These configurations have been set in favor of a model with a more balanced sensitivity and specificity values. In order to assess the predicitivity of the model, statistical analysis have been conducted in terms of accuracy, sensitivity, and specificity using cross validation routine as an internal evaluation, in addition to an external evaluation using an external test set. In this article, we name this model E.

Internal evaluation of the models

Accuracy, sensitivity, and specificity have been determined for the internal evaluation of each model using the SARpy program. For the internal validation, fivefold cross-validation routine was conducted for each model. In the five-fold cross validation the learning set is randomly partitioned into five equal sized subsets. For each iteration, a single subset of chemicals was retained as the validation data for testing the model, and the remaining subsets were used as training data. The cross validation process was repeated five times (the folds). The evaluation results of five iterations were then averaged to produce a single estimation. Accuracy, sensitivity, and specificity of the internal evaluation are assessed in addition to the Matthews correlation coefficient (MCC).

External evaluation of the models

The predictability of the models has been evaluated on two external test sets: the first external set is the dataset used as the learning set of the opposite model (e.g., for the R model we used ISSCAN-CGX dataset and vice versa), and the second dataset is a collection of 258 compounds collected from the eChemPortal inventory. Accuracy, sensitivity, specificity, and the MCC for the external evaluation are determined using SARpy. Although the external evaluation is considered the best mean for the assessment of the predictive ability of a (Q)SAR model,[43,44] the results of the external evaluation of any model are highly related to the relative similarity of the external evaluation set in relation to the learning set.

Table 1. R model internal and external validation for five different splits and the average of the model performance.

		1° split (59 active rules)	2° split (65 active rules)	3° split (61 active rules)	4° split (58 active rules)	5° split (57 active rules)	Average
Learning set (778 compounds)	Accuracy	71%	72%	71%	70%	71%	71%
	Sensitivity	75%	75%	71%	73%	70%	73%
	Specificity	65%	69%	71%	66%	72%	69%
Test set (337 compounds)	Accuracy	63%	60%	64%	65%	62%	63%
	Sensitivity	68%	58%	62%	67%	61%	63%
	Specificity	56%	63%	66%	61%	64%	62%

Results and discussions

R model

Each learning set produced its own model, which is a collection of active SAs with their likelihood ratios. The final model merging all sets of SAs consisted of 127 active SAs. Table 1 shows the predictive performance of five models developed based on five different splits of the ANTARES database. The performance of each model has been evaluated on its own learning set using cross-validation analysis. Further, an external evaluation using the corresponding test set is performed on each model. To have an overview of the statistical analysis of the performance of the models, we calculated the average of the predictive values of all the five models, and reported in Table 1 as well. The averages of accuracy, sensitivity, and specificity for the 778 compound internal cross-validation using five rule sets extracted from the ANTARES dataset were 71%, 73%, and 69%, respectively. The average of accuracy, sensitivity, and specificity for 337 compounds in the test set as an external validation of these models, were 63%, 63%, and 62%, respectively.

Using the R model, the results of cross-validation on the whole training set were 66% accuracy, 83% sensitivity, 48% specificity, and 0.34 the MCC (Table 2). Analysis of the external validation for the R model demonstrated that the concordance between experimental and predicted value on the ECHA dataset is higher than using the ISSCAN-CGX dataset. The accuracy of the R model on the ECHA dataset was 67%, compared to 58% of accuracy for the ISSCAN-CGX dataset. The complete list of these alerts are presented in the VEGA platform.

Table 2. R model and E model internal and external validation.

	R model (127active rules)			E model (43active rules)		
	Crossvalidation	External validation on ISSCAN and CGX data	External validation on ECHA data	Crossvalidation	External validation on ANTARES data	External validation on ECHA data
Accuracy	66%	58%	67%	73%	59%	64%
Sensitivity	83%	76%	62%	77%	77%	48%
Specificity	48%	40%	70%	62%	41%	72%
TP[a]	651/783	593/735	55/89	562/735	599/783	43/89
TN[b]	367/760	142/254	119/169	157/254	315/760	121/169
FP[c]	393/760	112/254	50/169	95/254	445/760	48/169
FN[d]	132/783	142/735	34/89	172/735	184/738	46/89
MCC[e]	0.34	0.35	0.31	0.36	0.19	0.20

[a]True positive; [b]True negative; [c]False positive; [d]False negative; [e]Matthews Correlation Coefficient.

E model

With the configuration set as mentioned above, SARpy extracted 43 active rules from the ISSCAN-CGX learning set. Analysis of the cross-validation for the E model demonstrated that the second model produced an accuracy of 73%, with a sensitivity of 77% and a specificity of 62% (Table 2). The MCC value for this analysis is 0.36. The accuracy values for the external evaluation of the E model on the ANTARES dataset and the ECHA database were 59% and 64%, respectively. Analysis of the external validations for the E model demonstrated that the model produced a higher sensitivity (77%) compared with the specificity (41%) of the R model. On the contrary, the specificity of the external evaluation on the chemicals from the ECHA database was higher (72%) compared to its sensitivity (48%) (Table 2). The complete list of the SAs present in this model is accessible through VEGA.

Analysis of the combination of the prediction results of the R and the E models

Another analyses has been done on the prediction results of the R model and the E model. In this new approach, we considered the final results as correctly predicted only in case both models have predicted them consistently. Table 3 summarizes the results of combining the R and E model external validation predictions on the chemicals from the ECHA database.

The results suggested that when both models are concordant on a negative prediction for a compound the reliability of the result is much higher than in case a positive prediction is done. We observe an improvement of the results compared to the use of the individual models, for accuracy (72%) and specificity (79%). In fact, combining the predictions of the two models the

MCC is increased to 0.37, compared to 0.31 for the R model and 0.20 for the E model. Only sensitivity is higher using the R model (62%). Thus, users may choose a solution or another depending if they prefer a conservative or a realistic assessment.

Table 3. The combination of the predictions of the R and E models on the ECHA external validation set.

	Combined model
TP[a]	33/89
TN[b]	96/169
FP[c]	25/169
FN[d]	24/89
Accuracy	72%
Sensitivity	58%
Specificity	79%
MCC[e]	0.37
Coverage	178/258

[a]True positive; [b]True negative; [c]False positive; [d]False negative; [e]Matthews correlation coefficient.

Fragments analysis

Comparison of the SAs in the R and E models

The SAs present in the R and E models have been compared and those that are in common between the two rule sets categorized into chemical classes and listed as follows. The SAs in the R model are presented with their ID number and written in order of their correspondence to the identical SAs in the E model.

1. Aromatic amine (R model: 6, 41, 36, 22, 10 / E model: 27, 31, 33, 38, 104)
2. Aromatic heterocyclic (R model: 12, 19, 2 / E model: 75, 108, 117)
3. Hydrazide (R model: 28, 27 / E model: 2, 50)
4. N-Nitroso (R model: 1 / E model: 8)
5. Phenyl-Hydrazine (R model: 32 / E model: 48)
6. α,β- Haloalkanes (R model: 25 / E model: 56)
7. Sulfite (R model: 8 / E model: 68)
8. Nitrogen Mustard like (R model: 11 / E model: 73)
9. Phosphonite (R model: 15 / E model: 98)

Categorization of the SAs in the R and E models

The SAs present in the models R and E are categorized from a chemical class point of view. The substructures within each category are presented with their ID number in their original rule set and are as follows:

Nitrogen containing substructures (Azo type):

1. Aromatic amine (R model: 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 33, 34, 35, 36,37, 38, 40, 42, 83, 104, 110, 113 / E model: 6, 10, 22, 31, 35, 36, 41, 42)
2. Aromatic heterocycles containing Nitrogen (R model: 74, 75, 80, 81, 83, 95,113, 122 / E model:12, 17, 43)
3. Azine (Hydrazine) (R model:46, 47, 49, 50, 51, 53, 54, 55, 101 / E model: 27,32)
4. Azide (Hydrazide) (R model: 2, 3, 44, 45, 52 / E model: 3, 28)
5. Nitrosamine (R model:4, 5, 7, 9, 10 / E model: not found (NF))
6. Nitrogen or sulfur mustard (R model: 72, 73, 115 / E model: 11, 34)
7. Aromatic methylamine (R model: 30, 34, 36 / E model: NF)
8. Aliphatic N-Nitroso (R model: 62, 63/ E model: NF)
9. Aromatic Nitro (R model: 90, 123 / E model: NF)
10. 1 aryl 2 monoalkyl hydrazine (R model: 48 / E model: NF)
11. Aziridine (R model:120 / E model: NF)
12. Aromatic hydroxylamine (R model: 32 / E model: NF)
13. Diazo (R model:92 / E model: NF)
14. Aromatic Azo (R model: 71 / E model: NF)
15. Aromatic Nitroso (R and E models: NF)

Other substructures:

1. (1,2, and 3 membered) Aromatic Heterocycles (R model: 74, 75, 80, 81, 83, 90, 95, 103, 108, 113, 117, 121, 122, 123 / E model: 2, 12, 17, 19, 43)
2. Aliphatic halide (R model: 57, 58, 59, 70, 125 / E model: 18, 25)
3. Heterocyclic Alkane (R model: 84, 105, 109, 120 / E model: 23)
4. Polycyclic aromatic systems (R model: 39, 43, 60, 61 / E model: 30)
5. Sulfonate bonded carbon (R model: 67, 68 / E model: 8)
6. Epoxide (R model: 105 / E model: 23)
7. B propiolactone (R model: 114 / E model: NF)

Not only SARpy was able to find the already known carcinogen substructures that were represented by the SAs of Kazius and colleagues, [45] but a number of SAs have been identified for the first time. Table 4 demonstrates the new identified SAs that have been classified into six chemical classes. The substructures within each category are listed with their ID number and are as follows:

1. Nitrosurea (R model: 12, 13, 14, 19 / E model: NF)
2. Nitrogen or sulfur mustard like (R model: 72, 115 / E model: 34)
3. Benzodioxole and Benzendiol (R model: 17, 18 / E model: 9)
4. Tertiary amine substituted by a Sulfur atom (E model: 24)
5. α,β-oxy and carboxy substitutions (R model: 20, 21, 76 / E model: NF)
6. α,β-haloalkanes (R model: 56, 69 / E model: 25)

7. Oximes (R model: 78 / E model: NF)

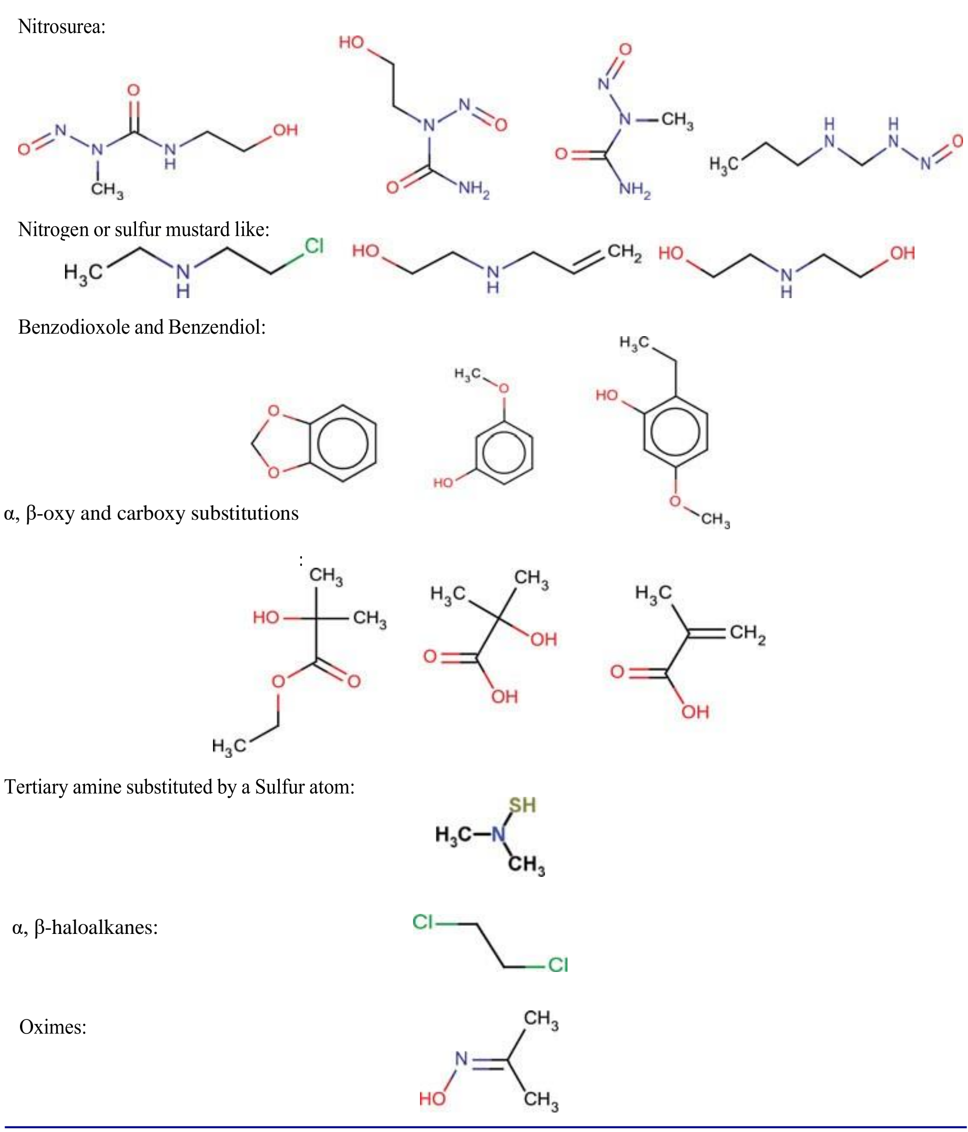
For the sake of example, we illustrated the chemicals from which the SA 24 (form the chemical class tertiary amine substituted by a Sulfur atom) in the E model has been extracted (Table 5). All the chemicals that contain the previously mentioned SA in the ISSCAN-CGX data set are carcinogenic.

Discussion

Automated extraction of SAs has been implemented by the statistically-based program SARpy on two learning sets. The ANTARES learning set collects rodent bioassay carcinogenicity data on 1543 chemicals, while ISSCAN-CGX database containing 986 chemicals takes into account human-based assessments and data retrieved from different assays. The predictive performance of the developed models were evaluated internally as well as using a 258 compound external validation dataset collected from the ECHA inventory. The two developed models for carcinogenicity have been implemented in the VEGA platform and are indeed freely available for end users.

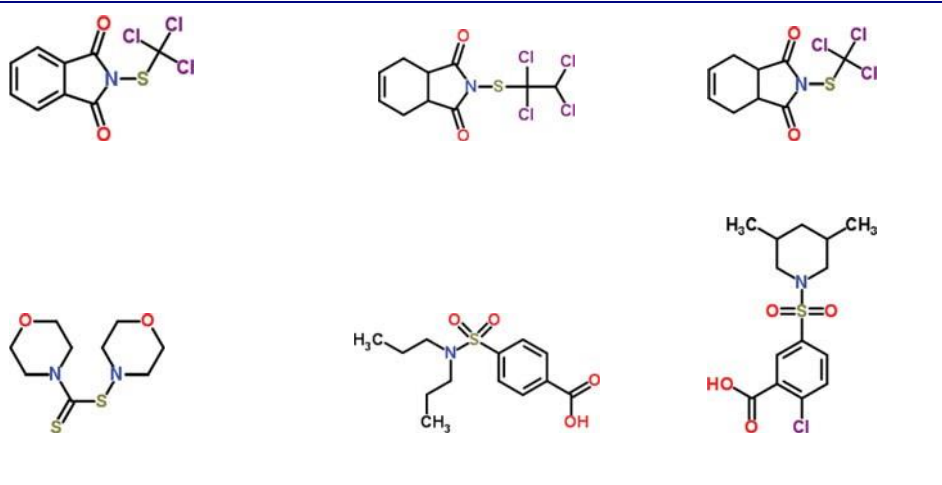
Recent progresses in data mining provide effective competence in the automated discovery of SAs associated with toxicological endpoints. An important contribution of the statistically based methods to the carcinogenicity field is identification of new SAs, which help us in refining the existing rule sets. While the most known carcinogenicity rule sets [23] are composed on the basis of human expert judgement, the SAs identified in our study are extracted in an unbiased manner by SARpy with no a priori knowledge about the MoA of the chemicals. This approach sheds light to the new clues about genotoxic and nongenotoxic SAs. Some primary analyses have been provided on the SA lists; chemical classes of the identified SAs have been evaluated; however, further study for the new SAs should be performed considering other collections of alerts [45]. SARpy SAs resulting from the current analysis on the ANTARES and ISSCAN-CGX data sets follow the SAs presented by Kazius and colleagues [46].

Table 4. New carcinogenic structural alerts identified by SARpy in the R and E models.



Further, the models are developed on the basis of two learning sets with different carcinogenicity data from the point of view of origin and provenance. Concerning the learning sets with substantially variant carcinogenicity data assessed within different properties, each set of the extracted SAs constituted a purpose-oriented model. The user may consider the results of the model with more realistic predictions or the one with more conservative assessments.

Table 5. Chemicals structures in the ISSCAN-CGX data set from which structural alert 24 has been extracted.



Generally, the best approach in making a conclusion to estimate the reliability of a prediction is combining evidence from different information sources such as (Q)SAR model predictions, *in vitro* and *in vivo* test results. This is reflected in the general trend of developing ensemble models and/or combining the output of different existing models. An example of the latter approach has been done on a similar endpoint, mutagenicity (Ames test), by the integration of the different models available on the VEGA platform [47]. The advantage of having the two presented models available on the VEGA platform, where other models for the same endpoint are available, is also the possibility of performing a similar activity to make a conclusion.

Finally, the results of the presented models will be exploited for the improvement of ToxRead (http://www.toxgate.eu), a recent platform that uses set of rules for different endpoints to filter and select similar compounds and assist the user in performing read-across studies [48,49] . Also, these rules can be compared and possibly explained considering reasoning about mechanisms, including adverse outcome pathways [50].

Funding

The research for this article was financially supported by the Life PROSIL project and Ministero della Salute of Italy/Istituto Superiore di Sanità (ISS) within the project "Messa a punto di strategie integrate di testing—basate su metodi alternativi—per l'identificazione di sostanze cancerogene per il REACH."

References

[1] Regulations R. No. 1272/2008 of the European Parliament and of the Council, on classification, labeling and packaging of substances and mixtures, amending and repealing directives 67/548/EEC and 1999/45/EC, and Amending Regulation (EC) No. 1907/2006, *Official J. Eur Union.* 2008:L353.

[2] Arcos JC Chemical Induction of Cancer: Modulation and Combination Effects. *An Inventory of the Many Factors Which Influence Carcinogenesis*.SpringerScience&BusinessMedia: Birkhäuser Boston; 1995.

[3] Woo Y, Lai D. Mechanisms of action of chemical carcinogens, and their role in StructureActivity Relationships (SAR) analysis and risk assessment. In: *Quantitative StructureActivity Relationship (QSAR) models of mutagens and carcinogens*. (R. Benigni, ed.) 2003:41– 80.

[4] Huff J. Long-term chemical carcinogenesis bioassays predict human cancer hazards: Issues, controversies, and uncertainties. *Ann N Y Acad Sci.* 1999;895(1):56–79.

[5] Tomatis L. Identification of carcinogenic agents and primary prevention of cancer. *Ann N Y Acad Sci.* 2006;1076(1):1–14.

[6] Doll R. The causes of cancer. *Revue d'épidémiologie et de santé publique.* 2001;49(2): 193.

[7] Ames BN, Gold LS. Chemical carcinogenesis: Too many rodent carcinogens. *Proceedings of the National Academy of Sciences*. 1990;87(19):7772–7776.

[8] Anisimov VN, Ukraintseva SV, Yashin AI. Cancer in rodents: Does it tell us about cancer in humans? *Nat Rev Cancer.* 2005;5(10):807–819.

[9] Knight A, Bailey J, Balcombe J. Animal carcinogenicity studies: 1. Poor human predictivity. *Altern Lab Anim*. 2006;34(1):19–27.

[10] Knight A, Bailey J, Balcombe J. Animal carcinogenicity studies: 2. Obstacles to extrapolation of data to humans. *Altern Lab Anim*. 2006;34(1):29–38.

[11] Ward JM. The two-year rodent carcinogenesis bioassay—Will it survive? *J Toxicol Pathol.* 2007;20(1):13–19.

[12] Ashby J, Paton D. The influence of chemical structure on the extent and sites of carcinogenesis for 522 rodent carcinogens and 55 different human carcinogen exposures. *Mutat Res, Fundam Mol Mech Mutagen.* 1993;286(1):3–74.

[13] Ward JM, Rat or mouse cancer bioassay—Or none of the above? *Toxicol Pathol.* 24(6):734– 735.

[14] Louekari K, Sihvonen K, Kuittinen M, Sømnes V. In vitro tests within the REACH information strategies. *Alternatives to laboratory animals. ATLA.* 2006;34(4):377–386.

[15] Combes R, Grindon C, Cronin M, Roberts DW Garrod JF. Integrated decision-tree testing strategies for mutagenicity and carcinogenicity with respect to the requirements of the EU REACH legislation. *Alternatives to laboratory animals. ATLA.* 2008;36:43–63.

[16] Wells MY, Williams ES. The transgenic mouse assay as an alternative test method for regulatory carcinogenicity studies—Implications for REACH. *Regul Toxicol Pharmacol.* 2009;53(2):150–155.

[17] Benigni R, Bossa C. Predictivity of QSAR. *J Chem Inf Model.* 2008;48(5):971–980.

[18] Benigni R, Bossa C. Predictivity and reliability of QSAR models: The case of mutagens and carcinogens. *Toxicol Mech Methods.* 2008;18(2–3):137–147.

[19] Miller J, Miller E. Origins of human cancer. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY; 1977:605–627.

[20] Miller EC, Miller JA. Searches for ultimate chemical carcinogens and their reactions with cellular macromolecules. *Cancer.* 1981;47(10):2327–2345.

[21] Ashby J, Tennant RW. Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the US NCI/NTP. *Mutat Res.* 1988;204(1):17–115.

[22] Benigni R, Bossa C, Tcheremenskaia O, Giuliani A. Alternatives to the carcinogenicity bioassay: In silico methods, and the in vitro and in vivo mutagenicity assays. *Expert Opin Drug Metab Toxicol.* 2010;6(7):809–819.

[23] Benigni R, Bossa C. Structure alerts for carcinogenicity, and the Salmonella assay system: A novel insight through the chemical relational databases technology. *Mutat Res.* 2008;659(3):248–261.

[24] Benigni R, Bossa C, Tcheremenskaia O. Nongenotoxic carcinogenicity of chemicals: Mechanisms of action and early recognition through a new set of structural alerts. *Chem Rev.* 2013;113(5):2940–2957.

[25] Toxtree, Ideaconsult Ltd, Sofia, Bulgaria. http://toxtree.sourceforge.net/.

[26] Rosenkranz H. Quantitative Structure-Activity Relationship (QSAR) models of chemical mutagens and carcinogens. In: Quantitative Structure-Activity Relationship (QSAR) models of chemical mutagens and carcinogens. Boca Raton, FL, CRC Press; 2003;175–206.

[27] OncoLogic, Environmental Protection Agency. http://www2.epa.gov/tsca-screeningtools/oncologictm-computer-system-evaluate-carcinogenic-potential-chemicals.
[28] Derek Nexus, Lhasa Limited, Leeds, UK. http://www.lhasalimited.org.
[29] Gini G, Ferrari T, Cattaneo D, Golbamaki Bakhtyari N, Manganaro A, Benfenati E. Automatic knowledge extraction from chemical structures: the case of mutagenicity prediction. *SAR QSAR Environ Res.* 2013;24(5):365–383.
[30] VEGA. Istituto di Ricerche Farmacologiche Mario Negri Milano. http://www.vega-qsar.eu.
[31] ANTARES project, Grant Agreement LIFE08 ENV/IT/00435. http://www.antares-life.eu/.
[32] CAESAR project, no. 022674—SSPI. http://www.caesar-project.eu/.
[33] Gold LS, The Carcinogenic Potency Project and Database (CPDB). University of California, Berkeley; Lawrence Berkeley, National Laboratory; National Library of Medicine's (NLM®); 2011. http://potency.berkeley.edu.
[34] Leadscope. http://www.leadscope.com/.
[35] Benigni R, Cecilia B. ISSCAN: Istituto Superiore di Sanita, "Chemical carcinogens: Structures and experimental data." https://w3.iss.it/site/BancaDatiCancerogeni/.
[36] Kirkland D, Aardema M, Henderson L, Müller L. Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens: I. Sensitivity, specificity and relative predictivity. *Mutat Res, Genet Toxicol Environ Mutagen.* 2005;584(1):1–256.
[37] istMolBase. Kode. http://chm.kode-solutions.net/products_istmolbase.php.
[38] InstantJChem. ChemAxon. https://www.chemaxon. com/ products/ instant-jchem-suite/ instant-jchem/.
[39] Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J Chem Inform Comp Sci.* 1988;28(1):31–36.
[40] eChemPortal. In: Organisation of Economic Co-Operation and Development. http://www.echemportal.org/echemportal/index?pageID=0&request_locale=en.
[41] Classification Labelling and Packaging inventory. In: European Chemicals Agency. http://echa.europa.eu/it/regulations/clp/cl-inventory.
[42] Ferrari T, Gini G, Bakhtyari NG, Benfenati E. Mining toxicity structural alerts from SMILES: A new way to derive Structure Activity Relationships. In: Computational Intelligence and Data Mining (CIDM), IEEE Symposium on: 11–15 April 2011: 120–127.
[43] Tropsha A, Gramatica P, Gombar VK. The importance of being earnest: Validation is the absolute essential for successful application and interpretation of QSPR models. *QSAR Comb Sci.* 2003;22(1):69–77.
[44] Perkins R, Fang H, Tong W, Welsh WJ. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology. *Environ Toxicol Chem.* 2003;22(8):1666–1679.
[45] Ceccaroli C, Pulliero A, Geretto M, Izzotti A. Molecular fingerprints of environmental carcinogens in human cancer. *J Environ Sci Health, Part C: Environ Carcinog Ecotoxicol Rev.* 2015;33(2):188–228.
[46] Kazius J, McGuire R, Bursi R. Derivation and validation of toxicophores for mutagenicity prediction. *J Med Chem.* 2005;48(1):312–320.
[47] Cassano A, Raitano G, Mombelli E, Fernández A, Cester J, Roncaglioni A, Benfenati E. Evaluation of QSAR Models for the prediction of Ames genotoxicity: A retrospective exercise on the chemical substances registered under the EU REACH regulation. *J Environ Sci Health, Part C: Environ Carcinog Ecotoxicol Rev.* 2014;32(3):273–298.
[48] Gini G, Franchi AM, Manganaro A, Golbamaki A, Benfenati E. ToxRead: A tool to assist in read across and its use to assess mutagenicity of chemicals. *SAR QSAR Environ Res.* 2014;25(12):999–1011.
[49] Benfenati E, Manganelli S, Giordano S, Raitano G, Manganaro A. Hierarchical rules for read-across and in silico models of mutagenicity. *J Environ Sci Health, Part C: Environ Carcinog Ecotoxicol Rev.* 2015;33(4):385–403.
[50] Benigni R, Battistelli CL, Bossa C, Giuliani A, Tcheremenskaia O. Alternative toxicity testing: Analyses on skin sensitization, toxcast Phases I and II, and carcinogenicity provide indications on how to model mechanisms linked to adverse outcome pathways. *J Environ Sci Health, Part C: Environ Carcinog Ecotoxicol Rev.* 2015;33(4):422–443.

Structural Alerts for Carcinogen Compounds (R Model)

Following, the list of the 127 rules for carcinogenicity, expressed as SMARTS strings:

SA 1: CN[N+]=O

SA 2: NNC=O

SA 3: CN(C=O)N=O

SA 4: CCCCCCN(C)N=O

SA 5: CCCN(CCC)N=O

SA 6: CNCCNN=O

SA 7: CNCCN(C)N=O

SA 8: CCNN=O

SA 9: CCCCCN(C)N=O

SA 10: CCCCN(C)N=O

SA 11: CC(O)CNN=O

SA 12: CN(N=O)C(=O)NCCO

SA 13: NC(=O)N(CCO)N=O

SA 14: CN(N=O)C(N)=O

SA 15: CCCNN=O

SA 16: O(c1ccccc1)c2ccccc2

SA 17: COc1cccc(O)c1

SA 18: CCc1ccc(OC)cc1O

SA 19: CCCNCNN=O

SA 20: CCOC(=O)C(C)(C)O

SA 21: CC(C)(O)C(O)=O

SA 22: Cc1ccccc1-c2ccc(N)cc2

SA 23: Nc1ccc(cc1)-c2ccc(N)cc2

SA 24: Nc1ccc(cc1)-c2ccccc2

SA 25: Nc1ccc(C=C)cc1

SA 26: Nc1cccc(c1)-c2ccccc2

SA 27: Cc1ccc(N)c(C)c1

SA 28: Cc1ccccc1N

SA 29: Nc1ccc(Cc2ccc(N)cc2)cc1

SA 30: CN(C)c1ccc(Cc2ccccc2)cc1

SA 31: Cc1ccc(N)cc1

SA 32: Cc1ccc(NO)cc1

SA 33: Cc1cccc(N)c1

SA 34: CNc1ccc(C=C)cc1

SA 35: Nc1ccc2ccccc2c1

SA 36: CNc1ccc(N)cc1

SA 37: Nc1ccccc1O

SA 38: COc1ccccc1N

SA 39: Oc1cccc2ccccc12

SA 40: Nc1ccc(O)c(N)c1

SA 41: CC(C)C(C)(O)CO

SA 42: Nc1cccc(c1)S(O)(=O)=O

SA 43: Cc1cccc2ccccc12

SA 44: NNCO

SA 45: CN(N)CO

SA 46: CC(O)CNN

SA 47: NNCCO

SA 48: NNc1ccccc1

SA 49: NNCC=C

SA 50: CCNN

SA 51: CCCNN

SA 52: CC(=O)NN

SA 53: CCCN(N)CCC

SA 54: CCCN(C)N

SA 55: NN

SA 56: ClCCCl

SA 57: CCl

SA 58: CCBr

SA 59: CBr

SA 60: c1ccc2cc3c(ccc4ccccc34)cc2c1

SA 61: c1ccc-2c(c1)-c3cccc4cccc-2c34

SA 62: CN=O

SA 63: N=O

SA 64: NO

SA 65: Oc1cccc(c1)-c2cccc(O)c2

SA 66: OS(O)(=O)=O

SA 67: COS(O)=O

SA 68: COS(=O)=O

SA 69: ClC1CCCC(Cl)C1Cl

SA 70: CC(Cl)CCCCCl

SA 71: c1ccc(cc1)N=Nc2ccccc2

SA 72: CCNCCCl

SA 73: ClCCNCCCl

SA 74: Cc1ccncn1

SA 75: c1cncnc1

SA 76: CC(=C)C(O)=O

SA 77: CC=C(C)CO

SA 78: CC(C)=NO

SA 79: CC(C)=N

SA 80: Cn1cncn1

SA 81: c1ncnn1

SA 82: COc1ccc(CC=C)cc1

SA 83: Nc1ncc2ncn(CCCCO)c2n1

SA 84: OCC1OC(CC1O)n2cnc3cncnc23

SA 85: CC1CCC=C(C)C1

SA 86: C1C=CCC=C1

SA 87: O=C(OCc1ccccc1)c2ccccc2

SA 88: CCOCc1ccccc1C

SA 89: C(=Cc1ccccc1)c2ccccc2

SA 90: [O-][N+](=O)c1ccco1

SA 91: CCNCC(C)=O

SA 92: N=[N+]

SA 93: Cc1ccc(cc1)S(O)(=O)=O

SA 94: O=C1c2ccccc2C(=O)c3ccccc13

SA 95: Cc1cccnc1

SA 96: CCCCC(O)CCCCC(O)CCC

SA 97: Clc1cccc(Cl)c1Cl

SA 98: COP=O

SA 99: CC(CN)c1ccccc1

SA 100: OCC#C

SA 101: NNCc1ccccc1

SA 102: C1CCc2ccccc2C1

SA 103: c1ccsc1

SA 104: Nc1ccccn1

SA 105: C1CO1

SA 106: CC(O)CCCC=O

SA 107: C[S]=O

SA 108: c1cscn1

SA 109: CC1COCO1

SA 110: Nc1ccc([S]c2ccccc2)cc1

SA 111: Cc1ccc(cc1)C(N)=O

SA 112: CN(C)P(N(C)C)N(C)C

SA 113: [N+]c1cncn1

SA 114: O=C1CCO1

SA 115: OCCNCC=C

SA 116: CCNCCCC(C)C

SA 117: c1ccoc1

SA 118: CCOC(N)=O

SA 119: C=CCCCCC=O

SA 120: C1CN1

SA 121: c1cc2ccccc2s1

SA 122: Cc1ncc[nH]1

SA 123: [O-][N+](=O)c1ccc(o1)-c2cscn2

SA 124: C#C

SA 125: CCF

SA 126: CN=[N+]

SA 127: CCCN=CN

Structural Alerts for Carcinogen Compounds (E model)

Following, the list of the 43 rules for carcinogenicity, expressed as SMARTS strings:

SA 1: O=NNCC

SA 2: c1occc1

SA 3: O=CN(N)C

SA 4: CCCN(CC)CC

SA 5: C1CC(=CC)CCC1

SA 6: Nc1ccc(cc1C)C

SA 7: NCCCN

SA 8: O=S(=O)(OC)

SA 9: c1ccc2OCOc2c1

SA 10: Nc1ncccc1

SA 11: N(CCCl)CCCl

SA 12: c1cn(cnc1)

SA 13: C=C(C=C)C

SA 14: O=NNC

SA 15: O=P(OC)

SA 16: O(c1ccc(cc1)CC=C)

SA 17: c1ncn(c1)C

SA 18: C(CCCC(CC)Cl)Cl

SA 19: c1ncsc1

SA 20: C=CCN

SA 21: O=Cc1ccccc1O

SA 22: O(c1ccc(cc1N))C

SA 23: O1CC1C

SA 24: SN(C)C

SA 25: C(CCl)Cl

SA 26: c1c(cc(cc1Cl)Cl)Cl

SA 27: NNCC

SA 28: O=CN(N)

SA 29: C(OC)C(C)C

SA 30: c1ccc2cc(ccc2c1)

SA 31: Nc1cccc(c1C)C

SA 32: NNc1ccccc1

SA 33: c1cc(ccc1C)Cl

SA 34: N(CCO)CCO

SA 35: Nc1ccc(cc1N)

SA 36: c1ccc(cc1N)C

SA 37: O(c1ccc(cc1)C)C

SA 38: C(c1ccccc1)CO

SA 39: C(=CCC)CC

SA 40: N(Cc1ccccc1)C

SA 41: Nc1ccc(cc1)C

SA 42: Nc1ccccc1

SA 43: n1cccc(c1)

Golbamaki et al., Drug Des 2017, 6:3

Drug Designing: Open Access

ISSN: 2169-0138

DOI: 10.4172/2169-0138.1000153

Research Article OMICS International

ToxDelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances

Azadi Golbamaki[1]*#, Alessio M Franchi[2#], Serena Manganelli[1], Alberto Manganaro[1,3], Giuseppina Gini[2] and Emilio Benfenati[1]

[1]IRCCS – Istituto di Ricerche Farmacologiche Mario Negri, Via Giuseppe La Masa, 19, 20156 Milan, Italy
[2]Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milan, Italy
[3]Kode s.r.l., Via Nino Pisano, 14, 56124 Pisa, Italy
[#]Contributed equally

Abstract

ToxDelta is a new tool for the evaluation of the toxicity of chemicals based on a modified version of the fmcs_R package. Two structurally similar molecules share a maximum common substructure (MCS). In order to evaluate if two similar molecules have different effects, we focused our attention on the molecular fragments which are not in the MCS. These parts may increase or decrease the value of the property. We considered a variation of the MCS concept of efficient relevance in toxicity assessment where the rings of molecules must not be broken. To assess the toxicity of the target chemical, ToxDelta extracts the MCS and delineates the remaining fragments. Each of these moieties represents a difference between two molecules and its relevance in the toxicity assessment is evaluated against a knowledge-based list of active and inactive fragments. ToxDelta considers the dissimilarities of the molecules in a read-across approach.

ToxDelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances

Keywords: Read-across; (Q)SAR, SAR; Maximum Common Substructure; Molecular graph; Toxicity assessment; Mutagenicity; Structural alert

Introduction

The European REACH legislation for industrial chemicals promotes the use of alternative methods, and explicitly mentions and regulates the use of read-across and quantitative structure-activity relationships (QSAR), jointly named non-testing methods (NTM). Often, one of the problems of QSAR models is their poor interpretability. Along with the assessment of the predictive power and statistical quality, the interpretability of the QSAR models is an important issue for the regulatory bodies. Furthermore, since read-across is related to the concept of similarity, it is closer to evidence and apparently easier to be accepted, although similarity cannot be univocally defined. The European Chemicals Agency (ECHA) published a document with the purpose to communicate the framework applied within the agency to evaluate the assessment done with read-across [1]. In 2014, it was reported that the most common and widely used NTM consisted in building categories and predicting properties by read-across. Up to 75% of the analysed dossiers contained read-across at least for one endpoint. The ECHA guidance on QSARs and grouping of chemicals introduces a flowchart [2] for the generation and use of non-testing data in the regulatory assessment of chemicals. This flowchart consists of a sequence of operations (eight steps), which starts with information collection and terminates with the final assessment exploiting the functionalities of a vast range of computational tools and databases. Depending on the chemical and property of interest these steps can be omitted or performed in a different order. In our new tool ToxDelta, we have addressed two of these steps: "Search for structural alerts" and "read-across". Even though read-across has been used much more than QSAR for registrations, it has been studied much less than QSAR. There are many open issues on the use of this approach. Read-across is typically subjective, and strongly relies

on the individual expert, the expert's background and experience, and is difficult to reproduce [3].

In order to solve the above mentioned problems of interpretation and to help the expert to get a documented, transparent and reproducible evaluation on the activity of the target compound, our group developed a new read-across tool: ToxRead [3,4]. This tool assists experts in the evaluation of the biological activity/toxicity of compounds, offering known elements affecting the activity within the same picture.

Recently, we published the results of an exercise on read-across. Participants made their assessment using the approach they preferred. The group of scientists who used ToxRead gave consistent assessment for the same chemical, while those who used other programs typically gave conflicting assessment [5]. This indicates that the subjectivity of the assessment may introduce a source of variability which may make read-across an unreliable strategy without a proper reproducible scheme.

Generally, the programs assisting the expert in read-across are based on similarity measuring software. Examples of these programs are ToxRead, the OECD QSAR toolbox [6], ToxMatch [7] and AMBIT [8]. VEGA, which is commonly used for QSAR, can be also used as a read-across tool, as it shows the similar compounds and in many of its models also the alerts [3]. Similarity is basically measured on the basis of the chemical structure. In some cases additional toxicological considerations are added. These programs just show the most similar compound(s) to the target substance, and the user can decide the biological activity/toxicity of the target compound on the basis of the activities of the similar compounds used as source compounds. Furthermore, some of these programs (e.g. ToxRead) provide the value predicted by the software.

***Corresponding author:** Azadi Golbamaki, IRCCS – Istituto di Ricerche Farmacologiche Mario Negri, Via Giuseppe La Masa, 19, 20156 Milan, Italy, Tel: +39 02 3901 4742/4595; E-mail: azadi.golbamaki@marionegri.it

Received September 05, 2017; **Accepted** October 04, 2017; **Published** October 10, 2017

Citation: Golbamaki A, Franchi AM, Manganelli S, Manganaro A, Gini G, et al. (2017) ToxDelta: A New Program to Assess How Dissimilarity Affects the Effect of Chemical Substances. Drug Des 6: 153. doi: 10.4172/2169-0138.1000153

Copyright: © 2017 Golbamaki A, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

In combinatorial chemistry, the use of similarity and diversity methods addresses the similarity property principle. This principle states that structurally similar molecules have similar biological activity [9]. This statement is questioned by various experiences with contradictory results [10]. In fact, structural similarity does not always imply similarity in either activity [11] or descriptors [12]. Minor modifications can make active molecules to lose their activities completely and vice versa. Intrinsically, the similarity concept includes the fact that the two molecules are different. Thus, the expert should evaluate not only how similar the two substances are, but also whether the differences trigger an opposite behaviour.

Many similarity measuring methods have been proposed to quantify the similarity between chemical compounds especially in drug discovery research. One of the most famous methods is the study of substructure and superstructure relationships of the chemicals. Two molecules may share some common properties due to their common substructure. This search strategy does not provide any quantitative similarity measurement. Hence, it is a very knowledge-based approach in which every substructure used in a query needs to be well defined. Structural descriptor-based methods are another commonly used structural similarity searching approach in which the similarity of the chemical compounds can be quantified. Structure similarity search does not require an exact match and the search results are ranked by scores. One of the important *structural-based* search methods is fingerprint [13]. In this method, the chemical structure is disclosed in a highdimensional space. Many models for predicting biological activities are based on the similarity coefficient provided by such methods, such as neural networks [14], fuzzy adaptive least squares [15] and inductive logic programming [16]. Structural descriptor-based methods are computationally simple, but they are unable to identify local similarities between structures.

Maximum common substructure (MCS) is an encouraging approach for similarity searching and biological activities predictions in chemoinformatics. The MCS is a problem of graph matching that involves 2D or 3D chemical structures of two chemicals and identifies the largest substructure present in both molecule structures. The MCS-based methods have all the advantages of the substructure and superstructure -based methods and in addition does not need an exact match procedure. Compared to structural descriptor–based methods, MCS provides a

similarity score for each comparison and can perform local similarity identification. MCS is a straightforward concept of determining similarities with a clear chemical meaning and is principally independent of the fingerprints. Several available MCS algorithms in the literature do not satisfy the graph representation of the chemical compounds. Barrow and Burstall in 1976 [17] used the MCS concept for the sub-graph isomorphism for the first time. After that, Cone et al. [18] introduced the use of MCS for similarity search for molecular comparison. The approach did not receive a notable consideration due to its complexity. Later, other *MCS-based* similarity search algorithms have been presented [19-21]. The concept of MCS in the molecule structures has been applied in different chemoinformatic concepts, such as classification models using the structural similarities [22], enrichment of chemical libraries [23] and clustering molecules with similar structural features [24]. The MCS search methods are mainly divided into “clique” [17,25] and “backtracking” [26,27]. The computational problem of finding all the largest complete sub-graph(s) (maximal clique) is called the clique problem. The clique problem is NP-complete, i.e., no polynomial time algorithm has been found to solve the general problem. However, many algorithms for computing cliques have been developed, both complete and approximate. The basic algorithm is due to Ullmann [28] who introduced backtracking to reduce the size of the search space. The MCS extraction algorithm of the FMCS R package is based on backtracking search method.

We considered a variation of the MCS concept of efficient relevance in toxicity assessment, where the rings of molecules must not be broken. We modified the MCS algorithm of the fmcs_R package [29] for finding the MCS between two given similar molecule graphs subject to this constraint. The similarity index is determined by the VEGA similarity indication which is described described by Floris et al. [30]. The new software, ToxDelta is a novel read-across tool developed to identify and extract the differences between the target and the reference compounds for the further evaluation of the biological activity/toxicity of the target molecule. These differences are depicted as molecular substructures. Their possible role in amplifying or reducing the activity/toxicity of a compound is queried in an *a* priori prepared data set of molecular structural alerts (SA).

We added the constraint of keeping the aromatic or aliphatic rings present in the target or the reference molecule complete during the process of the MCS extraction. Indeed, this decision is made due to the important role that rings play in the mutagenicity and carcinogenicity SAs. For example, polycyclic aromatic hydrocarbons (PAHs) that are composed of multiple aromatic rings are a class of mutagens. In addition, PAHs are linked to skin, lung, bladder, liver, and stomach cancers in confirmed animal models. The increasing number of aromatic rings in PAHs helps the metabolic activation to reactivate diol epoxide intermediates and consequently their binding to DNA [31]. In addition, the mutagenicity of the aliphatic epoxides has been determined by the Ames test [32]. Historically, a very effective list of the SAs has been created and revised by Ashby in 1985 and 1988, respectively [33,34]. The Ashby's well-known poly-carcinogen list contains aromatic nitro groups, aromatic azo groups, aromatic rings N-oxides, aromatic mono- and di-alkylamino groups, aromatic amines and aliphatic and aromatic epoxides. The extended SAs list according to Kazius et al. [35] contains groups of specific aromatic nitro and amine, aliphatic halide, polycylic aromatic system and other SA with aromatic or aliphatic rings. Also, the Benigni's [36] list includes an important number of forms of aromatic and aliphatic rings.

In our study, we combine a substructure identification tool with a tool for the assessment of the related fragments which are not in common, in order to evaluate the toxicity of the two chemical compounds under examination.

At present, ToxDelta performs only mutagenicity assessment of the chemical substances. The mutagenicity SAs collection of this tool is extracted from the Ames test results. Ames test is the gold standard for initial examination for detecting chemically induced gene mutations of new chemicals and drugs. Bruce Ames created the Ames assay in the 1970s [37]. The assay's sensitivity towards many types of mutagens has been improved over the years [38,39]. Specific distinct mutations in the histidine and tryptophan synthetic pathways of *Salmonella typhimurium* and *Escherichia coli* have been created respectively, that result in the requirement for an exogenous supply of those amino acids for growth. Using genetically bacterial strains, the Ames test produces a high rate of inter-laboratory reproducibility (85%-90%) [40]. This assay has been proved to be the most predictive *in vitro* assay for rodent and human carcinogenicity

[37,41]. Additionally, the Ames test results have been demonstrated to be in agreement with rodent carcinogenicity or *in vivo* genetic toxicity about 65% [42].

Materials and Method

Database of active and inactive structural alerts

In a previous study, a very sophisticated collection of mutagenicity SAs has been created and implemented in ToxRead for the read-across mutagenicity assessment [4]. This set of rules associated to bacterial mutagenicity has been identified and extracted by analyzing more than 6000 chemicals from different chemical classes. A set of rules related to both mutagenicity and lack of mutagenicity were found. These SAs have been sorted in a hierarchy of rules and used to identify the active or inactive mutagenic substructures present in the target compounds. The hierarchical order of the SAs makes it possible to identify first the exact rule that matches the target molecule and then other, more generic ones, which may match with the target molecule. Besides rules for mutagenicity and non-mutagenicity, the identified potential rules include exceptions and modulators of activity. These rules can be also used to predict mutagenicity concerning the influence of each SA found in the molecule. Accuracy and p-value are two statistical characterizations which are assigned to each SA of the mutagenicity list; these values show the accuracy of the SA based on the number of chemicals in the original training set containing the SA, and the prevalence of one of the categories: Mutagenic or non-mutagenic. These SAs are those implemented within ToxRead. In the case of the module for mutagenicity, there are about 800 SAs each with a high level of detail such as accuracy and statistical significance.

The MCS algorithm: ToxDelta advances ToxRead for it supports the reasoning based on the differences, as well as similarity of molecules. The degree of similarity and dissimilarity between pairs of molecules is computed from their structures. Molecular structures can be encoded in several computer formats which basically contain the topological information about the structure, as well as other chemical information such as atom charges, aromaticity, etc. Among several available formats, we relied on SMILES strings [43] and structure data formats (SDF).

The algorithm proposed in the fMCS_R package [44] performs MCS computation via a novel backtracking algorithm by incrementally computing a search tree of correspondences between nodes of the two graphs under investigation. Each node in this tree is a set of atom correspondences while leafs are the connected sub-graphs we are looking for; the deepest leafs are the MCSs.

The aromatic and non-aromatic rings as structural properties of molecules and their role in the biological activities of the molecules are important issues. Indeed, among the identified mutagenic and carcinogenic SA, aliphatic and aromatic rings play an important role. In this regard, an important number of forms of rings are established [36]. For this reason, we applied the constraint that the rings present in the input molecular graphs must be retained by the MCS by adding an additional check in the backtracking algorithm of the fmcs_R library, as opposed to the original R package (fmcs_R). In other words, we decided to maintain all the rings entire and not break these rings during the process of the MCS extraction. Our modification to the original code of this library consists mainly in adding a restriction during the process of the atom selection for the MCS. For each atom belonging to a ring, this restriction checks if the atom resides in an equivalent ring in the target and the reference molecule. Since the fmcs_R algorithm does not consider rings as such, it may break some rings, i.e. if it is necessary it selects only a subset of atoms in a ring. This leads to a significant loss of structural information and consequently the implication of the extracted MCS which is meant to be equal for both molecules may differ for each compound. We can finally extract the structural differences between the two compounds under investigation: we overlap each graph with the MCS and highlight all the sub-branches not in the MCS (Figure 1).

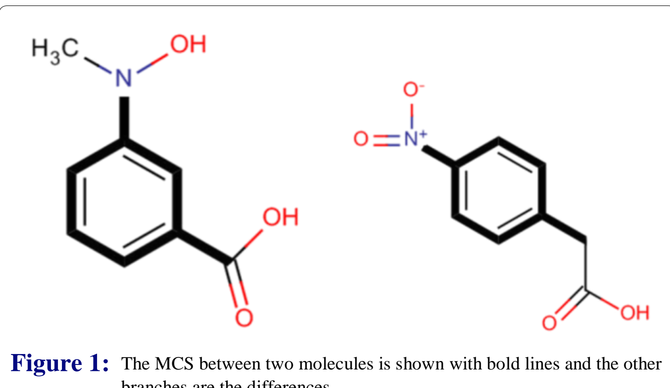


Figure 1: The MCS between two molecules is shown with bold lines and the other branches are the differences.

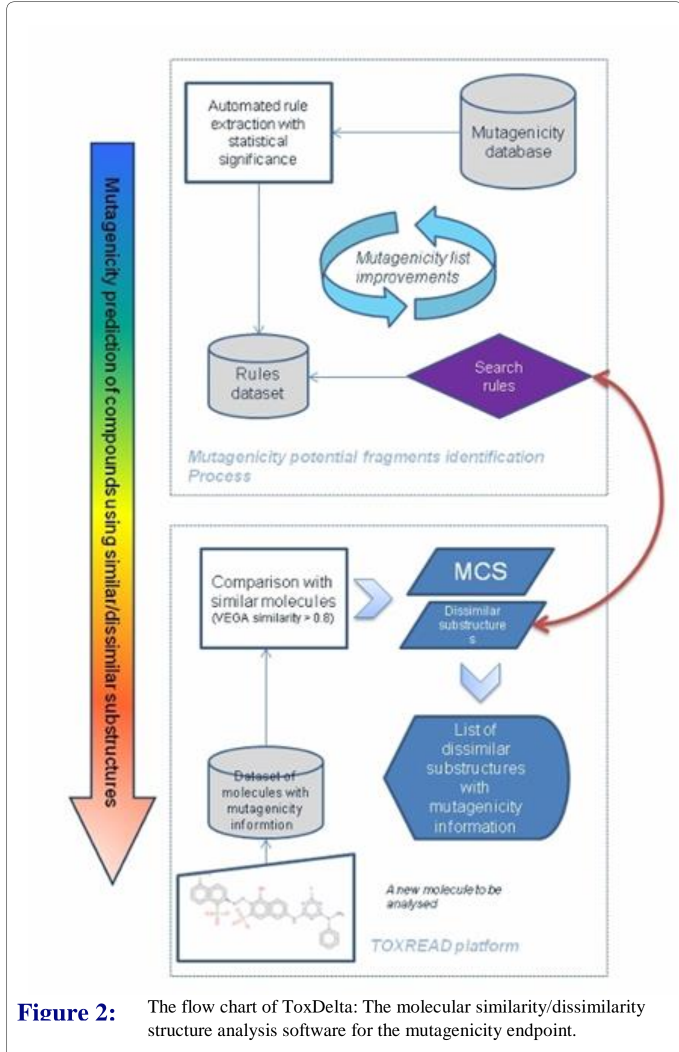


Figure 2: The flow chart of ToxDelta: The molecular similarity/dissimilarity structure analysis software for the mutagenicity endpoint.

ToxDelta Implementation

The user can evaluate the evidences obtained by ToxDelta and make a decision regarding the toxicity of a compound under evaluation, in a weight of evidence approach. Figure 2 shows the flow chart of the implementation of the ToxDelta program for the mutagenicity endpoint. The new tool relies on ToxRead for the evaluation of the degree of similarity between similar

compounds. The similarity algorithm has been described elsewhere [30]. A stand-alone version of ToxDelta is accessible on the VEGA home page (https://www. vegahub.eu/). ToxDelta will be implemented inside ToxRead and the dissimilar substructures will be computed between the target molecule and any source molecule selected by the user. ToxRead associates the most similar molecules present in its data base to the target molecule, pointing out the mutagenic (or non-mutagenic) fragment(s) as toxicity rules present in both the target and the similar chemical compounds. ToxRead identifies the mutagenic or non-mutagenic SAs in common between the target and the source chemicals. Thus, these SAs belong by definition to the MCS of the pair of compounds under investigation. At this point, the integration of ToxDelta inside ToxRead will allow further investigation of the pair of compounds, identifying the dissimilar moieties and providing the most similar SAs for each of them in the collection of the known SAs. To obtain a conceivable result, the structure of the target and the source molecules in the comparison need to be sufficiently similar. If the structures of the molecules compared by ToxDelta do not share a significant MCS, the dissimilarities may not be interpretable to an acceptable level. In other words, whenever the structures of two molecules are strongly dissimilar, the user may not expect a significant MCS. In this regard the VEGA chemical similarity index [30] is used as a screening before applying the MCS approach.

Provided that the identified dissimilar fragment in the target molecule is an SA along with the assigned active or inactive toxicity effect information, there are three possible scenarios that may help the user to move in a certain direction for toxicity decision-making. These three scenarios about the dissimilar fragment found in the target molecule are as following:

1. The SA is an active fragment with strong evidence which increases the effect;
2. The SA is an inactive fragment with strong evidence which decreases the effect;
3. The SA is a fragment without any relevant impact on the effect.

In case 1 and 2 the SA is more likely to modulate the effect of the whole molecule, while in case 3 the SA is a neutral fragment and does not have an impact on the modulation of the effect. Nevertheless, the software provides documentation on the SAs of case 3, which indicates the existence of a certain fragment with no impact on the effect. Documentation is an important factor in the acceptance of the read-across results. This whole list of SAs is used by ToxDelta to

assess whether the fragments resulting from the subtraction of the MCS from the molecule are associated to an increased or decreased or neutral effect.

The program returns as output all the possible MCSs of the same length (the number of the atoms is equal in all the MCSs) extracted from two molecules of interest. The user can choose one of the MCSs found and evaluate the dissimilarities calculated based on the chosen MCS. The different fragments present in both molecules, are the result of the subtraction of the MCS and the target or source molecules.

Evaluation of ToxDelta

In order to evaluate the new tool and to investigate its dissimilarity approach theory, we performed tens of in-house tests while studying and developing the proposed methodology. We selected two pairs of molecules with known mutagenicity (Ames test) experimental value as case studies, to show how our approach works and how it could be useful in the toxicity assessment. Even though there is no similarity threshold determined by this tool, for the molecules selected as case studies, we chose two pairs of molecules, case 1 and case 2, with a cut-off value of 0.7 for the VEGA similarity index [30]. The results provided by ToxDelta for the molecules with a small MCS may not have a significant interpretation. In both cases, we chose two compounds with different toxicity activity (one mutagenic and the other one non-mutagenic), as this scenario represents exactly the type of situation in which ToxDelta can provide useful insight. To check whether the structural differences between these molecules have a significant role in their toxicity or non-toxicity property, we assume that one of the molecules in each pair is the target molecule and the other one is the source molecule. We selected two pairs of derivatives from two relevant pharmaceutical classes: benzodiazepines and androstane derivatives. We chose diazepam, first came on the market as Valium, a benzodiazepine drug typically producing a calming effect. It is commonly used to treat anxiety, alcohol withdrawal syndrome, benzodiazepine withdrawal syndrome, muscle spasms, seizures, trouble sleeping, and restless legs syndrome. Flunitrazepam, known as Rohypnol, is a benzodiazepine derivative that can cause anterograde amnesia; its importation has been banned by the U.S. Government (https://chem.nlm.nih.gov/ chemidplus/name/flunitrazepam). The similarity VEGA index value between these two benzodiazepines is 0.87. Despite this, they

exhibit different toxicological profiles: Indeed Diazepam is experimentally non-mutagenic while flunitrazepam is mutagenic. As second case study, we provided two androstane derivatives: mepitiostane and a structural analogue, cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-. Metpitiostane is an antineoplastic agent inhibiting the expansion of estrogen-stimulated cancers by a competitive inhibition mechanism for the estrogen receptor (https:// pubchem.ncbi.nlm.nih.gov/compound/mepitiostane#section=Pharm acology-and-Biochemistry). The similarity index value between these two chemicals is 0.77. We processed the selected molecules using ToxDelta to explain the discrepancy between mutagenic activities for each pair. The results of the ToxDelta tool are discussed in the results section.

Results

The new ToxDelta software uses the structures of the two chemicals to be compared as input. The two substances are introduced as SMILES [43]. The MCS is the common part present in both molecules and it is shown (Table 1). This MCS is usually a large part of the molecules to be assessed. Indeed, the application of ToxDelta is useful for substances that are structurally similar. The MCS typically, even if implicitly, represents the driving force in the read-across procedure. This is the logical process which identifies the analogies among substances. In this scheme, ToxDelta does not contradict but complements the read-across conceptual strategy. The risk of the read-across strategy is to miss the differences between two molecules. The similarity should not erase the possible opposed behaviour of the two similar compounds. But how to avoid the error of ignoring factors which may provoke opposite behaviour? ToxDelta wants to address this issue. It carefully identifies the differences and the related toxicological meaning. The theoretical basis is closely related to the SA paradigm. Thus, ToxDelta complements the ToxRead software, which exploits all the SAs of the target compound. Beyond this global assessment, done by ToxRead, it is useful to apply ToxDelta for a closer look at the two substances (i.e. the target and the reference compounds), in particular, when they may have opposite toxicological properties. Indeed, it should be reminded that ToxRead predicts the

toxicological property of the target compound, and thus the predicted value of the target compound may be the opposite of the experimental value of the similar compound.

Table 1: The two case studies: Case study 1) target molecule: Diazepam, source molecule: Flunitrazepam; Case study 2) target molecule: cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-, source molecule: mepitiostane and the results of ToxDelta: Maximum common substructure and dissimilar fragments.

		Molecules	MCS*	Dissimilar fragments
Case study 1	Target	Cl, N, N, O, H_3C	N, N, O, H_3C	Cl CH_3 (Aromatic ring)
	Source	O^-, N^+, O, F, N, N, O, H_3C		F CH_3 (Aromatic ring); O^- N^+ O CH_3 (Aromatic ring)
Case study 2	Target	O, O, CH_3, H_3C, H_3C, S	H_3C, H_3C	O CH_3 (Aliphatic ring); O CH_3; (Aliphatic ring) S
	Source	Br, O, O, S, S, O, O, CH_3, CH_3, H_3C, CH_3, H_3C		(Aliphatic ring) O O S S O O (Aliphatic ring); H_3C H_3C CH_3 H_3C; Br CH_3 (Aliphatic ring)

* Maximum common substructure

Case study 1: Benzodiazepine derivatives

Source molecule 1: Flunitrazepam

Systematic name: 1,3-dihydro-5-(o-fluorophenyl)-1-methyl-7nitro-2H-1,4-benzodiazepin-2-one

SMILES: c12C(=NCC(=O)N(c1ccc(c2)[N+](=O)[O-])C)c1c(cccc1)F

Experimental activity: Mutagenic in Ames test [45]

CAS number: 1622-62-4

Target molecule 1: Diazepam

Systematic name: 1-methyl-5-phenyl-7-chloro-1,3-dihydro-2H1,4-benzodiazepin-2-one

SMILES: O=C1N(c3ccc(cc3(C(=NC1)c2ccccc2))Cl)C

Experimental activity: Non-mutagenic in Ames test [45]

CAS number: 439-14-5

ToxDelta identifies “1-methyl-5-phenyl-2,3-dihydro1H-1,4-benzodiazepin-2-one” as MCS shared by these two chemicals (Table 1). ToxDelta also extracts three fragments of dissimilarities: the nitro group, the fluorine and chlorine atoms, each linked to an aromatic carbon. Diazepam lacks the first two fragments, which are present in Flunitrazepam. The nitroaromatic moiety matches two ToxRead SAs for mutagenicity both referring to the generic nitroaromatic ring; the Benigni–Bossa alert does not include chemicals with orthodistribution and with a sulphonic group on the nitroaromatic ring. This leads to a slight difference in the accuracies of these fragments, which are respectively 85% and 87%. ToxDelta identifies also the fluorine and chlorine atoms linked to aromatic carbons as dissimilarity fragments between the two molecules. These moieties do not match any rule for Ames mutagenicity included in the ToxRead “libraries” of SAs [3,4]. As a conclusion, ToxDelta immediately reports as a key difference the presence of the nitroaromatic fragment, which is at the basis of the different mutagenicity value of the two substances.

Case Study 2: Androstane derivatives

Source molecule 2: cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-

Systematic name: Cholestan-6-one, 3-bromo-, cyclic 1,2-ethanediyl mercaptole, S,S,S',S'-tetraoxide, (3-beta,5-alpha)-

SMILES: O=S5(=O)(CCS(=O)(=O)C35(CC1C4CCC(C(C)
CCCC(C)C)C4(C)(CCC1C2(C)(CCC(CC23)Br)))

CAS number: 133331-34-7

Experimental activity: Mutagenic in Ames test [45]

Target molecule 2: mepitiostane

Systematic name: 5-alpha-Androstane, 2-alpha,3-alpha-epithio-17beta-(1-methoxycyclopentyloxy)-

SMILES: O(C)C6(OC2CCC3C4CCC1CC5C(CC1(C)
C4(CCC23(C)))S5)(CCCC6)

CAS number: 21362-69-6

Experimental activity: Non-mutagenic in Ames test [45]

ToxDelta identifies the androstane tetracyclic system as MCS shared by these two chemicals and extracts five fragments of dissimilarity (Table 1). Three of these are aliphatic rings: the thiirane, 1,1-dimethoxycyclopentane, and 1,3-Dithiolane 1,1,3,3-tetraoxide rings and two are aliphatic chains: the 2-methylheptyl group and a bromine atom, both linked to an aliphatic carbon ring. The cyclic moieties and the alkyl carbon chain do not match any rule potentially responsible for mutagenic/non-mutagenic activity listed in the ToxRead software. Conversely, the bromine atom linked to an aliphatic carbon ring corresponds to two ToxRead SAs both referring to bromo-/haloethyl moieties with different levels of specificity and a prevalence of mutagenic activity of 71% and 67%, respectively. These rules, which are present in the source molecule but not in the target chemical, give a first indication of different toxicological profiles for these chemicals.

Evaluating the two case studies, it is important to notice that sometimes the identified dissimilar fragment is not an entire SA. In many cases the fragment of dissimilarity is a fraction of the whole rule (an already existing rule in the rule set) and the rest of the SA appears in the MCS. This issue is completely solved by the ToxRead software. In fact, the dissimilarities examination of ToxDelta takes place after the visualization of the results of ToxRead, when the user has

already observed all the existing SAs that are matched with the target molecule and are in common between the target molecule and a set of structurally similar molecules.

Discussion and Conclusion

ToxDelta is a new tool for read-across concept not aimed at substituting other tools, but to complement them. It has been designed to match certain features of ToxRead, but it can also be used alone. It is important to underline that ToxDelta addresses differences between two molecules, and per se it does not address the overall toxic property of the molecule, while this aim may be accomplished by other tools, like ToxRead, covering the assessment of the target molecule. The main advantage of ToxDelta to the other read-across programs is its focus on dissimilarities in addition to the similarities and the resembling properties between structurally similar compounds. It exploits the adverse effects that these dissimilar fragments may trigger in the biological activities or properties of the chemical substances.

ToxDelta provides a further insight by analysing the modulations of the effects which are expected in relation to the presence of the additional fragments in one of the two molecules under evaluation. Compared to other tools for read-across, ToxDelta is more "local", and this fact makes it an ideal tool to evaluate the effect of the metabolites and the impurities related to a compound having at hand the experimental values for the parent compound. Two important fields in which this issue can be applied are impurities in pharmaceutics and pesticides. The Food and Drug Administration (FDA) has provided a guideline for industry about the mutagenicity of the pharmaceutical impurities [46] that describes a practical framework for identification and control of the identified mutagenic impurities in order to limit potential carcinogenic risk. Another appropriate field of application for this tool is in pesticides assessment. The European Food Safety Authority (EFSA) has addressed the possible use of *in silico* methods for the evaluation of the effects of metabolites of pesticides [47]. ToxDelta may represent an ideal tool for pesticides, biocides and pharmaceutical compounds; because in these cases the experimental property values of the parent compound is requested by the relative regulations and ToxDelta can provide this information. Thus, ToxDelta may be particularly useful in those cases where data for the parent compound are available, and the user is interested not in the absolute effect of the related compound, but the possible increase of effect in an impurity product. For instance, if the toxicity level of the impurities is similar to the parent compound, this fact does not affect the way the substance with the impurity should be handled

and regulated. Conversely, if the impurity represents an increased hazard, this may be a serious issue. To overcome this kind of problems, local tools that deal with measuring the relative increase or decrease of the effects are probably more accurate than absolute *de novo* predictions. ToxDelta aims to address an important issue associated with read-across. Although the use of read-across approaches is widespread, the acceptance of the dossiers using read-across is not straightforward. Detailed documentation has to be provided by the expert. One of the main sources of scepticism on the assessment of read-across is that there are two (or more) substances under consideration, the target compound, lacking of data, and the reference compound, which is assumed to represent the properties of the target compound. So far the existing software for read-across have focused on the assessment of similarity between the target and the source compounds, with the idea that the higher the similarity is, the higher is the likelihood that the properties of the two compounds will be similar. However, authorities often argue that even minor modifications of the chemical structure may provoke a dramatic change in the property value. To complement the existing software addressing similarity, we focused our attention on the differences between two compounds, introducing ToxDelta.

It is noticeable that unlike other read-across programs, the SAs within ToxRead and ToxDelta do not exclusively contain active fragments, but also inactive fragments. This advantage allows exploring positive and negative modulations of the effect, and recognizing whether any relevant impact is expected. These SAs are associated to statistical characterizations, based on the number of chemicals containing the fragment, and on the prevalence of one of the categories: toxic or non-toxic. As a result, the user has both, the evidence that a certain fragment is associated to a certain effect and the statistics related to the prevalence of active or inactive compounds containing that SA. ToxRead provides all the data available on mutagenicity and BCF endpoints, and enables the user to access the available knowledge in a read-across approach. This software with its genuine graphical user interface organizes different groups of similar molecules and allows the user to move in different levels of reasoning. ToxDelta nicely complements ToxRead, offering additional focus on all the fragments which may affect the toxicity.

Currently, a beta version of ToxDelta is freely available on the VEGA platform (https://www.vegahub.eu/) and the toxicity endpoint for which this tool can be used is mutagenicity. Other endpoints will be added to the software in the next future.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. Azadi Golbamaki and Alessio Mauro Franchi contributed equally.

Acknowledgement

This research was supported by the PROSIL project (LIFE12 ENV/IT/000154) and the EU-TOXRISK project. The EU-ToxRisk project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 681002.

References

1. (2015) ECHA. Read-across assessment framework (RAAF).
2. (2008) ECHA M. Guidance on information requirements and chemical safety assessment.
3. Franchi GG, Manganaro AM, Golbamaki A, Benfenati AE (2014) ToxRead: A tool to assist in read across and its use to assess mutagenicity of chemicals. SAR QSAR Environ Res 25: 999-1011.
4. Manganelli BE, Giordano SS, Manganaro ARG (2015) Hierarchical rules for read-across and in silico models of mutagenicity. J Environ Sci Health 33: 385403.
5. Belli MBE, Casimiro EBT, Fernandez ACJ, Honma MGG, Knauf RKM, et al (2016) Results of a round-robin exercise on read-across. SAR QSAR Environ Res 27: 371-384.
6. http://www.qsartoolbox.org.
7. https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_ toxicology/qsar_tools/toxmatch.
8. http://cefic-lri.org/lri_toolbox/ambit/.
9. Willett P, Barnard JM, Downs GM (1998) Chemical similarity searching. J Chem Informat Comput Sci 38: 983-996.
10. Doucet JP, Panaye A (1998) 3D Structural information: From property prediction to substructure recognition with neural networks. SAR QSAR Environ Res 8: 249-272.
11. Martin YC, Kofron JL, Traphagen LM (2002) Do structurally similar molecules have similar biological activity? J Med Chem 45: 4350-4358.
12. Kubinyi H (2002) Chemical similarity and biological activities. J Braz Chem Soc 13: 717-726.
13. Chen X, Reynolds CH (2002) Performance of similarity measures in 2D fragment-based similarity searching: Comparison of structural descriptors and similarity coefficients. J Chem Inf Comput Sci 42: 1407-1414.
14. Burden FR (1996) Using artificial neural networks to predict biological activity from simple molecular structural considerations. Quantitative Structure-Activity Relationships 15: 7-11.
15. Moriguchi I, Hirono S, Liu Q, Nakagome I (1992) Fuzzy adaptive least squares and its application to structure-activity studies. Mol Inform 11: 325-331.
16. King RD, Srinivasan A (1997) The discovery of indicator variables for Qsar using inductive logic programming. J Comput Aided Mol Des 11: 571-580.
17. Barrow HG, Burstall RM (1976) Subgraph isomorphism, matching relational structures and maximal cliques. Inf Process Lett 4: 83-84.

18. Cone MM, Venkataraghavan R, McLafferty FW (1977) Computer-Aided interpretation of mass spectra. Molecular structure comparison program for the identification of maximal common substructures. J Am Chem Soc 99*:* 7668-7671.

19. Hagadone TR (1992) Molecular substructure similarity searching: Efficient retrieval in two-dimensional structure databases. J Chem Inf Comput Sci 32: 515-521.

20. Raymond JW, Gardiner EJ, Willett P (2002) Heuristics for similarity searching of chemical graphs using a maximum common edge sub graph algorithm. J Chem Inf Comput Sci 42*:* 305-316.

21. Raymond JW, Gardiner EJ, Willett P (2002) Rascal: Calculation of graph similarity using maximum common edge sub graphs. Comput Jl 45: 631-644.

22. Cuissart B, Touffet F, Crémilleux B, Bureau R, Rault S (2002) The maximum common substructure as a molecular depiction in a supervised classification context: Experiments in quantitative structure/biodegradability relationships. J Chem Inf Comput Sci 42*:*1043-1052.

23. Lisurek M, Rupp B, Wichard J, Neuenschwander M, von Kries JP, et al. (2010) Design of chemical libraries with potentially bioactive molecules applying a maximum common substructure concept. Mol Divers 14*:* 401-408.

24. Stahl M, Mauser H (2005) Database clustering with a combination of fingerprint and maximum common substructure methods. J Chem Inf Model 45: 542–548.

25. Raymond JW, Willett P (2002) Maximum common sub graph isomorphism algorithms for the matching of chemical structures. J Comput Aided Mol Design 16: 521-533.

26. Crandell CW, Smith DH (1983) Computer-assisted examination of compounds for common three-dimensional substructures. J Chem Inf Comput Sci 23:186197.

27. Xu J (1996) GMA: A generic match algorithm for structural homomorphism, isomorphism and maximal common substructure match and its applications. J Chem Inf Comput Sci 36:25-34.

28. Ullmann JR (1976) An algorithm for sub graph isomorphism. JACM 23: 31-42.

29. Wang Y, Backman TW, Horan K, Girke T (2013) fmcsR: Mismatch tolerant maximum common substructure searching in R. Bioinformatics 29: 2792-2794.

30. Floris M, Manganaro A, Nicolotti O, Medda R, Mangiatordi GF, et al (2014) Generalizable definition of chemical similarity for read-across. J Cheminform 6: 39.

31. Boström CE, Gerde P, Hanberg A, Jernström B, Johansson C, et al. Cancer risk assessment, indicators, and guidelines for polycyclic aromatic hydrocarbons in the ambient air. Environ Health Perspect 110*:* 451.

32. Wade DR, Airy SC, Sinsheimer JE (1978) Mutagenicity of aliphatic epoxides. Mutat Res Genet Toxicol 58: 217-223.

33. Ashby J (1985) Fundamental structural alerts to potential carcinogenicity or non-carcinogenicity. Environ Mutagen 7: 919-921.

34. Ashby J, Tennant RW (1988) Chemical structure, Salmonella mutagenicity and extent of carcinogenicity as indicators of genotoxic carcinogenesis among 222 chemicals tested in rodents by the U.S. NCI/NTP. Mutat Res 204*:* 17-115.

35. Kazius J, McGuire R, Bursi R (2005) Derivation and validation of toxicophores for mutagenicity prediction. J Med Chem 48: 312-320.

36. Benigni R (2008) The Benigni/bossa rulebase for mutagenicity and carcinogenicity – A module of toxtree. JRC Sci Tech Rep EUR 23241 1-78.

37. Ames BN, Durston WE, Yamasaki E, Lee FD (1973) Carcinogens are mutagens: A simple test combining liver homogenates for activation and bacteria for detection. Proc Natl Acad Sci U S A 70: 2281-2285.

38. Maron DM, Ames BN (1983) Revised methods for the salmonella mutagenicity test. Mutat Res Environ Mutagen Relat Subjects 113*:* 173-215.

39. Wilcox P, Naidoo A, Wedd DJ, Gatehouse DG (1990) Comparison of *Salmonella typhimurium* TA102 with *Escherichia coli* WP2 tester strains. Mutagenesis 5: 285-292.

40. Piegorsch WW, Zeiger E (1991) Measuring intra-assay agreement for the Ames Salmonella assay. In: Statistical Methods in Toxicology, Hothorn L (edtr,), Springer Berlin Heidelberg, Springer, pp: 35-41.
41. (2011) Guideline, IHT. Guidance on genotoxicity testing and data interpretation for pharmaceuticals intended for human use S2 (R1). In: International conference on harmonization of technical requirements for registration of pharmaceuticals for human use, ICH Expert Working Group, pp: 1-25.
42. Kirkland D, Aardema M, Henderson L, Müller L (2005) Evaluation of the ability of a battery of three in vitro genotoxicity tests to discriminate rodent carcinogens and non-carcinogens I. sensitivity, specificity and relative predictivity. Mutat Res 584: 1-256.
43. Weininger D (1988) SMILES A chemical language and information system. Introduction to methodology and encoding rules. J Chem Inf Comput Sci 28: 31-36.
44. Cao Y, Jiang T, Girke T (2008) A maximum common substructure-based algorithm for searching and predicting drug-like compounds. Bioinformatics 24: 366-374.
45. Hansen K, Mika S, Schroeter T, Sutter A, ter Laak A, et al. (2009) Benchmark data set for *in silico* prediction of ames mutagenicity. J Chem Inf Model 49: 2077-2081.
46. (2014) Guideline, IHT. Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk. In: International conference on harmonization of technical requirements for registration of pharmaceuticals for human use (ICH): Geneva.
47. (2005) EFSA, ER. 396/2005 of the European parliament and of the council of 23 February 2005 on maximum residue levels of pesticides in or on food and feed of plant and animal origin and amending Council Directive 91/414/EEC.