# [P8] Role of chemicals in the onset of Diabetes Mellitus Type I and II

Jacques Ehret[1], Michael Sattler[2,3], Gilles Marcou[4], Karl-Werner Schramm[1,5]

[1]*Molecular EXposomics (MEX), Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), Ingolstädter Landstr.1, 85764 Neuherberg, Germany.*
[2]*Chair of biomolecular NMR-Spectroscopy, Technische Universität München, Department of Chemistry, Lichtenbergstr. 4, 85748 Garching, Germany*
[3]*Institute of Structural Biology, Helmholtz Zentrum München - German Research Center for Environmental Health (GmbH), Ingolstädter Landstr.1, 85764 Neuherberg, Germany.*
[4]*Laboratoire d'infochimie, Université de Strasbourg, Faculté de Chimie, 2 rue Blaise Pascal, Strasbourg, France.*
[5]*Department für Biowissenschaften, Technische Universität München, Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt, Weihenstephaner Steig 23, 85350 Freising, Germany.*

The incidence of diabetes increases despite the increasing research investment on the subject. It is recognized that genetic predispositions and lifestyle choices are important risk factors in the development of the disease. Yet, there is growing scientific evidence that environmental contaminants contribute to the development of diabetes. Therefore, human exposure to environmental chemicals deserves more consideration. (Anti-)Diabetogenic chemicals have an effect on diabetes (onset, insulin resistance, glucose tolerance, pancreatic ß-cells), glycemia, lipidemia (cholesterol, triglycerides, free fatty acids) and body weight.

We built a database using expert analyses coupled with text mining methods. Quantitative Structure/Activity Relationship (QSAR) and Quantitative Description/Activity Relationship (QDAR) analysis were used to create predictive models. Targeted endpoints were the overall activity towards diabetes and diabetogenicity of chemicals. Descriptors used for QDAR analysis were keywords occurrence based on a library related to diabetes developed in-house. E-state, CDK, QNPR, and Chemaxon descriptors were used for QSAR analysis. Partial Least Squares, Artificial Neural Network, Decision Tree, Random Forest, and Support Vector Machine were used to calculate models.

The database is diversely populated by 2455 molecules, including 570 drugs, 518 pesticides, 93 cosmetics, and 73 food additives. The balanced accuracy of prediction models was between 60% and 89% depending on the method used, and models applied on an external test set of 42 molecules predicted diabetogenicity with a precision of 0.78 (E-state, J-48) and 0.70 (Chemaxon, J-48).

Our approach allows us to identify chemicals that can constitute potential risk factors in the development of diabetes (such as pesticides, fungicides, alimentary additives, anesthetics). We can also prioritize chemicals that may potentially lower the effects of diabetes (i.e. alimentary components or drugs). Finally, we identified correlated and associated terms for chemicals linked to diabetes.