

# Using statistical methods to predict the course of disease

April 13 2022, by Daniel Meierhans

---



At the ETH AI Center, Alexander Marx tackles medical questions in data science. The purpose is to bring together theory and practice. Credit: ETH Zurich / Nicola Pitaro

Data contains much more than just the information on the surface. With statistics, deeper cause-and-effect relationships can be brought to light. This is what Alexander Marx is researching as a Fellow at the ETH AI Center using artificial intelligence. One of his goals is to be able to make predictions regarding diabetes in children.

In diabetics, hypoglycemia usually doesn't occur randomly, just as a stock market price doesn't crash for no reason. That means both are also predictable, at least theoretically. In practice, however, such forecasts have so far succeeded only in the rarest of cases. But if Alexander Marx's project is successful, that will change for children with type 1 diabetes. "We're working on [predictive models](#) that can detect early on if there is a risk of hypoglycemia during the night," explains the ETH AI Center Fellow, adding: "When children engage in vigorous physical activity during the day, their blood glucose levels can drop below a critical threshold while they sleep. With a reliable forecasting model, this risk could be avoided."

## **Bringing cause-and-effect networks to light**

Marx is exploring this hypothesis as part of Julia Vogt's Medical Data Science Group. "I come from more of a theoretical background and have worked mostly with artificially generated data. The AI Center's purpose is to bring together theory and practice, which I find exciting. I now have to make my theoretical concepts work with real data."

Marx acquired his academic credentials at Saarland University in Saarbrücken, Germany. After completing a Master's degree in bioinformatics, he stayed on there to write his [doctoral thesis](#) at the Max Planck Institute for Informatics. His thesis examined causal discovery—[statistical methods](#) that can be used to create causal graphs from observational data, which make cause-and-effect networks visible.

## **Deriving predictions from correlations**

One way to apply these methods is to use survey data to identify all factors that are suspected of having an effect on a particular variable. A general example would be how a person's income depends on their age,

place of residence, gender, education, marital status or number of children. Based on the correlations found, predictions can then be made for individuals who were not surveyed. Marx clarifies that to do this, it's not even necessary to define the entire dependency chains; it's enough to elicit the smallest set of factors required to make a prediction.

## **From synthetic data to clinical reality**

With the help of [artificial intelligence](#) based on simulated data, Marx used these methods to study how the activities of about 500 selected genes in a human cell are related. Ideally, these methods can be scaled up in the future to include all of a cell's 25,000 or so genes. Such computer analyses of gene networks would easily and quickly provide biological and [medical research](#) with a comprehensive understanding of the processes that take place in a cell. Achieving this through laboratory experiments would require enormous effort, as the scientists would have to switch off each gene individually using genetic engineering tools and then measure how this affects the activity of all the other genes.

For the projects Marx is tackling at the AI Center, he needs to take causal discovery methods to a new level of complexity. Instead of using full [observational data](#) sets or synthetic data, as with gene expression, he now works with real data from [clinical practice](#). This makes the task markedly more difficult, as he soon found out: "In reality, individual pieces of information, measurements, or entire data sets are often missing, and how the data is collected also always differs from hospital to hospital and sometimes even from physician to physician."

## **Eliminating irrelevant correlations**

The clinical data that Marx analyzes for his prediction model in collaboration with physicians at the University Children's Hospital Basel

(UKBB) includes time series of pulse rate and blood glucose levels, as well as information on physical activities, caloric intake, insulin injections and sleep quality. It's then a matter of filtering the data to exclude any correlations from the model that are not related to the research question.

If the forecasts for a treating physician are to be robust and comprehensible, the number of factors must be kept as low as possible. Either way, it is too early to predict whether the model will be successful in practice: "With our project, we are venturing into areas that we haven't mastered yet with the methods available."

## **Nature, mountains and climbing in community**

In any case, the young researcher has got off to a successful start in Zurich. "When I first came here in the autumn, I felt at home very quickly. The city is very beautiful and the mountains are really close by," he says. A passionate rock climber, Marx particularly likes being so close to nature, especially the mountains. "Climbing lets me switch off and focus my attention completely on the grips. There's also the climbing community—I like to do things together with other people." In Saarbrücken, he was far away from the mountains and so mainly did bouldering indoors. Now that he's based in Zurich, he's looking forward to being able to visit alpine terrain more often.

## **Extraordinarily international and interdisciplinary**

Marx likes the AI Center just as much as he likes the city of Zurich and its surroundings: "The Center is extraordinarily international. There's also a great variety of subject areas. It's impressive and inspiring to be able to have peer-to-peer discussions with leading authorities from different scientific disciplines as a normal part of the daily routine."

The AI Center's interdisciplinary character is not limited to social interaction, however. Alongside bioinformatician Julia Vogt, Marx has a co-mentor in Peter Bühlmann, who specializes in high-dimensional statistics. These can be used to examine data sets in which many attributes are associated with each object. This includes the diabetes data that Marx analyzes. In addition, there is also an established collaboration with the Biomedical Informatics Group led by Gunnar Rätsch, who conducts research at the interface of machine learning and bioinformatics.

## **Learning from different data sources**

Marx himself is active in multiple subject areas. He has another project in which he explores what's known as multimodal learning. Here, the goal is to find commonalities in data from different sources. For example, he combines results from [positron emission tomography](#) (PET), which produces a 3D visualization of anomalies in tissue metabolism, with results from X-ray-based computed tomography (CT), which reveals density anomalies in layers of tissue.

A combination of the analysis of the two imaging methods, automated by machine learning, could drive major advances in tumor diagnostics. The vision is an AI system that finds the commonalities in the two data sets and applies them to issue reliable diagnoses and prognoses.

## **First experience as a lecturer**

At the moment, Marx is looking forward to his first lecture course, which he will give this coming Summer Semester together with colleagues from the AI Center. "I've always enjoyed working with students, and the Master's students at ETH are at a very advanced level. The discussions always generate input that I hadn't considered myself,"

Marx says. His fellowship thus allows him not just to hone his scientific skills, but also to gain initial experience as a lecturer.

Marx doesn't yet want to venture any concrete predictions about his own future, saying: "After my first experiences here, I'm confident that my time at ETH will prepare me excellently for both career options—academia and industry."

Provided by ETH Zurich

Citation: Using statistical methods to predict the course of disease (2022, April 13) retrieved 11 May 2024 from <https://medicalxpress.com/news/2022-04-statistical-methods-disease.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.