

Machine learning advances help to tackle cancer

August 5 2014

A new technique has been developed at the Institute of Food Research, which is helping to understand how epigenetic marks relate to the risk of developing cancer.

Epigenetic marks are changes to our genome, that don't alter the genes themselves but affect whether they are turned on or off. They have been associated with increased risk of developing diseases, including cancers, and are affected by diet and other environmental factors.

For this reason, epigenetics has been well studied, to map the places in the genome where the marks occur and to use this information to better understand the onset of disease. But there are different types of epigenetic marks, and they occur at thousands of sites across the genome, so understanding their significance relies on the use of mathematical techniques.

One such technique is machine learning. This involves looking for patterns in a set of training data, and using this to generate models that predict independent data.

Now, [machine learning](#) has been improved by Dr Thomas Wilhelm of the Institute of Food Research, which is strategically funded by the Biotechnology and Biological Sciences Research Council. Instead of developing one model from the training data, his technique involves developing hundreds of diverse models, and applying these to independent, unseen data, and seeing which models work best in their

ability to predict outcomes. This avoids 'overfitting' of a model to a specific training data set.

The new technique can be applied to many different situations, but Dr Wilhelm applied it to epigenetic data on cervical cancer. Millions of women are screened for cervical cancer in the UK each year, with over 3,000 being diagnosed with the condition. Worldwide it causes 400,000 deaths annually. The [human papilloma virus](#) (HPV) is the major cause of cervical cancer, but not every woman infected with it goes on to develop cancer. Epigenetics, and in particular DNA methylation markers, are associated with the onset of cancer. Dr Wilhelm used publicly-available case-control data of DNA methylation in women at various stages of developing cervical cancer to look for patterns that indicated a predisposition to the condition.

"We saw clear patterns of DNA methylation markers that predicted the development of [cervical cancer](#)," said Dr Wilhelm. "Intriguingly, the patterns still predict development of the condition even in women who hadn't been infected with HPV."

The new method significantly outperforms previous attempts to analyse patterns in the data. And there is potential for improvement through using larger data sets and looking at more epigenetic markers from the human genome.

Dr Wilhelm has already started a new collaborative project to develop a clinical test for triage of HPV infected women for further treatment. Two more projects about risk prediction of bowel and prostate cancer are being developed.

The new method is also opening up new lines of investigation in our understanding of diseases. One finding from the study seems to indicate that epigenetic [patterns](#) are actually linked to a predisposition to HPV

infection itself. Dr Wilhelm is now looking to collaborate with virologists to follow up this new theory, and uncover any more evidence of a link between epigenetics and virus susceptibility.

More information: "Phenotype prediction based on genome-wide DNA methylation data," Thomas Wilhelm, *BMC Bioinformatics* 2014, 15:193 [DOI: 10.1186/1471-2105-15-193](https://doi.org/10.1186/1471-2105-15-193)

Provided by Institute of Food Research

Citation: Machine learning advances help to tackle cancer (2014, August 5) retrieved 26 April 2024 from <https://medicalxpress.com/news/2014-08-machine-advances-tackle-cancer.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.