



Data Saves Lives case study detailed template

Downloaded from <https://datasaveslives.eu>

Version 1.0, October 2019

Copyright DataSavesLives

Research study title

Big data used to create a computerised animation for the early identification of cancer disease associations

A novel computer programme to help doctors and scientists to better understand which other diseases are likely to occur in patients with cancer.

Why the research was conducted

What disease or treatment of knowledge gap was studied?

This research was undertaken to improve medical and research understanding of which additional diseases are most likely to occur in patients with cancer. The work focused on patients with the nine most commonly occurring cancers. The researchers first looked at the diseases most often occurring before each patient developed the cancer, in case any of those diseases actually increased the risk of the cancer occurring. They then looked at diseases occurring after the cancer was diagnosed, in case the presence of the cancer increased a patient's risk of certain other diseases. Their aim was to develop a way of helping clinicians and researchers in the future to discover which other diseases might be associated with cancer, in a case some diseases increase the risk of others occurring in the same patient. The term co-morbidity used when a patient has more than one disease at the same time, which is relatively common, especially as people get older.

Why was this area of research needed?

Cancer has a huge impact on patients. It is also a major burden for health care systems across the world. Prevention and early detection play important roles in reducing this impact and burden. As an example, previous research has shown that women with type 2 diabetes have a higher risk of later developing breast cancer. Women with diabetes could therefore be offered more frequent breast cancer screening to help detect any breast cancer signs early. There might be many other connections between diseases and cancer that are not yet known, and so there might be further opportunities to help patients and health systems to detect early any signs of a type of cancer, if the risks are better understood. It is similarly important to know what other diseases a patient might be at risk of developing after they have had a cancer diagnosed and treated, which could be screened for and picked up early.

The data on multiple disease associations is complex to interpret. Computer programmes are needed to help scientists to better see these associations, by analysing large quantities of patient data and displaying the results to scientists in helpful ways.

When and by whom the research was conducted

Which organisation(s) conducted the research?

The College of Medical Science and Technology, Taipei Medical University, Taiwan

Which organisation(s) funded the research?

Several Taiwan ministries and universities provided research grants, including the Ministry of Science and Technology and the Ministry of Health and Welfare

When was the research undertaken (which years)?

This study was published in 2016
It used health data covering the 3 years from 2000 to 2002

The data that was used for the research

Which data sources were used?

Out-patient data on 782 million patients collected by the Taiwan National Health Insurance database

Which countries were involved in providing the data?

Taiwan

Which kinds of patient and health conditions were included?

The data included outpatient visits, dental visits, hospitalisations, medications prescribed, medications refilled, laboratory and imaging examinations, and procedure codes

What kinds of health data, from which data sources?

Data routinely stored and held by the Taiwan National Health Insurance database

How were the data collected or obtained, or accessed remotely?

The researchers obtained permission to use an extracted copy of the health insurance database for this research, covering the three years 2000-2002.

What was the legal basis for using the data?

There is no legal basis described in the research paper. It may not have been required since the data was anonymised.

How was participant identity protected, and the data kept securely?

All patient data was anonymised.

What kind of analysis was undertaken on the data?

The patient records in the database were first grouped and analysed according to each patient's age and gender. The computer programme was set up to detect diseases occurring in any patient within three years of each other, and to display these disease associations in ways that would help a scientist to determine the strongest disease connections. This data was accumulated for millions of patients, to highlight the disease associations that occur most often, and have not just coincided by chance.

This collection of disease associations was used to create a visual image that can be rotated or zoomed in to enable very detailed inspection.

The potential societal benefits that are hoped for from the research

What type scientific result was obtained?

An animated visualisation of cancer-disease associations was developed using a large medical dataset, called Cancer Associations Map Animation (CAMA)

What kind of innovation was developed or validated?

This has the potential to identify multiple diseases that are associated with cancer, by using existing medical data.

In which countries will the outcome be available, as findings, a product or service?

The results of this research have been openly published. The CAMA computer system is also openly available to researchers, worldwide.

Which health systems or other ecosystems will benefit from using it?

Clinicians will be able to identify co-morbidities / disease associations with cancer. The research community will have additional reference data for other kinds of research as well.

Which kinds of patient will benefit from this innovation?

Patients with cancers could benefit from this development. Clinicians can be more easily aware of which other diseases a patient might be at risk from, and seek to detect those early or possibly to them.

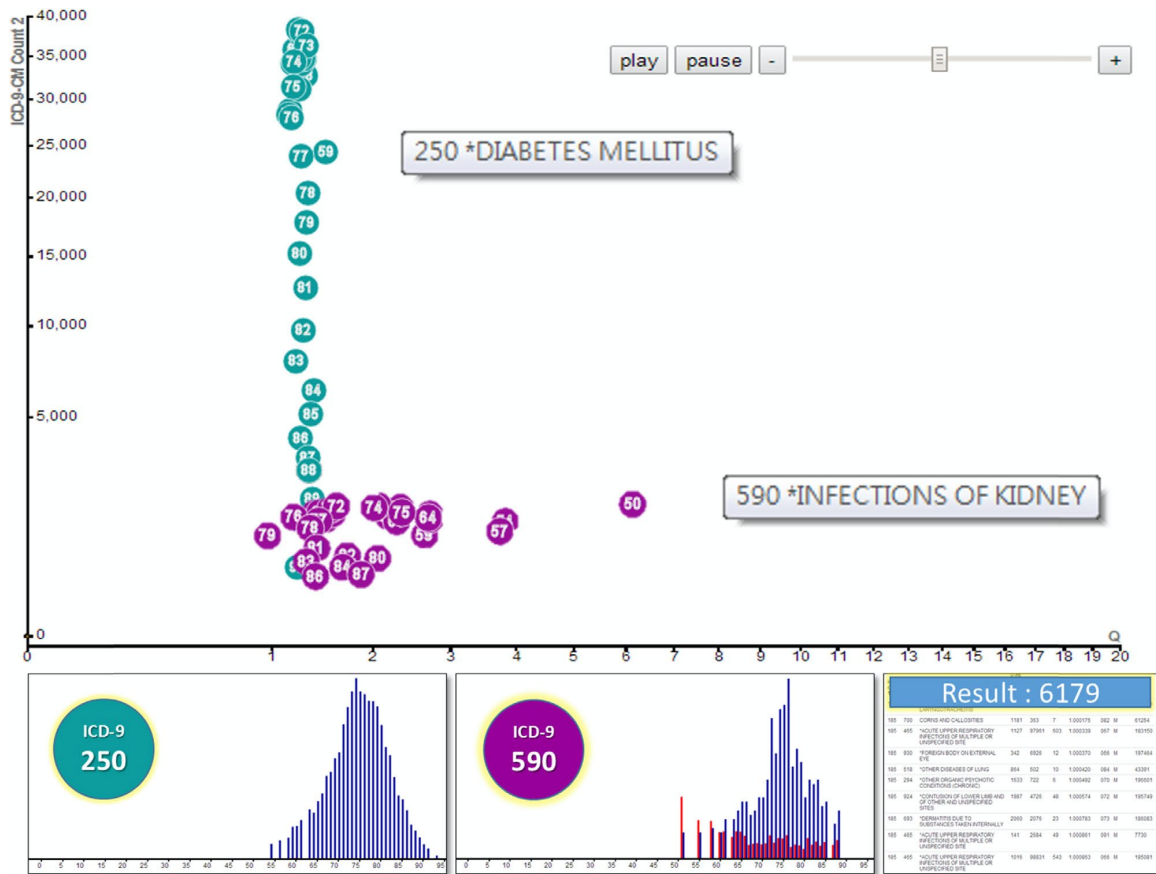
Additional information (if possible)

A link or reference to any more detailed findings, or the name of the intended/actual product or service

Iqbal, U. et al. (2016) Cancer-disease associations: a visualization and animation through medical big data. *Comput. Methods Programs Biomed.* 127, 44–51

A picture, graph, image or other visual that may liven up the web page, but is not too technical

Figure 3 from the paper is copied below.



Screenshot of the animation that would be used by researchers of clinicians showing the relationship of diabetes and infections of the kidney