Data science key to unlocking the community spread of COVID-19

July 1 2020

Credit: Pixabay/CC0 Public Domain

Medical scientists at the University of St Andrews have joined forces with data scientists Blue Hat Associates to further medical research and understanding about the spread of COVID-19 within the community.

How COVID-19 spreads within the community and individual households remains uncertain as current data is based on people who receive medical care.

A COVID-19 tracker app allows people to share data on the members of their household including symptoms.

This research, one of three research projects awarded to St Andrews as part of the Scottish Government's COVID-19 Rapid Research program, aims to extract key information from households' reports of COVID-19 and link these with medical records to get a better idea of the true number of people with the disease.

The project will look to see if people who stay at home with COVID-19 have different symptoms to those who need medical help or are admitted to hospital.

Using data being collected in the recently launched c19track.org website, this crowd-sourced data will help researchers understand the profile of the pandemic within households and within communities.

Key advantages of this tracking includes:

- Capturing data and updates on an entire household or neighborhood
- Tracking the disease in children
- Recording all symptoms of illness
- Simple technology allowing those without smartphones to record their data

St Andrews Medical School intend to use the data to research the outcomes for patients recovering in a home setting against those who seek GP support.

The data collected will be made available to research institutions and those signing up will be able to see the profile of their local area. Those signing up are agreeing to share their anonymous data with other organizations, and can opt in to share personal data for St Andrews medical research.

Professor Colin McCowan from the School of Medicine at St Andrews said: "One of the big unanswered questions with COVID-19 is how many people actually have caught it. We know about the people who contact the NHS but not those who have stayed at home self-isolating. This work will help us identify that group of people and allow us to examine if there are differences between them and the people who have contacted the NHS. This will give us a more accurate picture of how many people have COVID-19 and also help in planning how best we look after them."

Blue Hat founder, Tim Palmer, said: "As experts in data and analytics, we saw a lack of breadth in the data being collected, focussing on a narrow set of symptoms of those within a medical environment. We developed a crowd sourced data tracker in March to catch data from families who may only have minor symptoms and are delighted St Andrews will be using the data to aid the research."

What is the purpose of the website?

In March a group of data scientists launched c19track.org. At this point no one knows exactly how many people have been infected with COVID-19. There are many people with the virus who are invisible in the official numbers or only have mild symptoms. Without testing it is hard to know, so the aim is to help record peoples' symptoms to get a better snapshot of where the infections are and how many people really have the disease.

Who is Blue Hat?

Blue Hat is a team of data scientists who want to help, coming together to gather the missing information and provide it to the scientists who can help stop this virus. Blue Hat believes you can help by self-reporting online using this website, which will help understand more about the virus in our communities.

What is it trying to capture?

The website looks to build up a picture of the health for all members of a household, and for as many households in a geographic location as possible. This will provide crucial information and permit analysis at a postcode and local level of both the spread and current existence of the COVID-19 virus.

The data will help scientists to understand the profile of the disease at a local level, not a national or even county-wide level, but at a level which will be more relevant for each household, their local school and local community.

Is it any different to the others already out there?

Other symptom tracking apps are available, but Blue Hat's website allows an entire household's data to be captured on a simple web page allowing the scientists to track children and vulnerable people who could not fill in the forms for themselves. Occasional updates of this data are requested.

Unlike a number of apps which look to track a person's COVID-19 symptoms in considerable detail, asking a significant number of specific health questions, the Blue Hat website approaches pandemic data

gathering from the perspective of the community effect of the disease.

The aim is to build up a picture of the spread of the virus within households and within communities. This is something personal app trackers do not do, and is something that Blue Hat believes is going to become crucial in understanding the spread of the pandemic, the real level of infection in a community, and the potential for any local relaxation of quarantine in due course.

How does this link to the NHS contract tracking app?

Blue Hat's tracking works in an entirely different way and for a different purpose. The website is asking households to record contact information, creating a sample of a different cross section of the population without access to the latest technology.

What personal information will it hold?

The information collected by the website is simple, limited, and focused on the health of each member of a household in relation to COVID-19 symptoms. It includes:

- email address to communicate
- postcode to identify geographic location
- names of people of household
- health of each member of the household
- IP address for data security management

Isn't it becoming too late to capture this information?

It is clear that in a number of regions and communities there is a growing instance of suspected COVID cases, and tracking this information is

going to be increasingly important.

It is likely that there will be great benefit for researchers, for purposes such as immunization, to have accurate information and this can only be collected at the time a community is experiencing the pandemic. Surveys undertaken after the event are likely to be of limited use or have sufficient detail to be able to make critical judgements from.

Will my neighbors be able to see if I have had COVID-19 through the website?

Data will not be shared at a level where it will be possible to identify an individual household. Data will be anonymised and reporting is done on groups of households. The size of the area disclosed to a registered user will depend on the uptake in the area, but will not permit any sole households to be identified within a postcode.

How do I know my details will be safe?

The website has been properly registered with the information commissioners and will operate within all the regulations of the UK/EU in terms of data protection and GDPR. All information will remain within the EU and held securely on a globally recognized cloud platform.

Only aggregated and anonymised data will be shared—no personal information will be shared unless specifically requested and then only for the purposes of medical research. The underlying data will be made available free of charge to academic researchers working to understand the virus and helping to prevent future outbreaks.

Who has developed the website?

The c19track website has been developed by a group of experts in high security data systems who are used to working in government and banking technology software. The software team working on c19track are all affiliated with Blue Hat Associates, a London-based software development business.

What will St Andrews do with the information?,

Subject to approval from the individuals, St Andrews will combine the tracker data sourced in the community with NHS medical data.

Is anyone going to make money from the COVID-19 data?

No, the purpose of the website is to help medical researchers and academics better understand the spread of COVID-19, and for contributors to see the profile of their local area. The data is being made available to these groups for free, and individual data will only ever be shared for medical purposes and only if explicit consent has been given.

  More information: Help paint a more accurate picture of COVID-19 in your community: c19track.org

Provided by University of St Andrews