

# Teaching computers to read health records is helping fight COVID-19

19 October 2020, by James Teo, Richard Dobson



Doctors write information in free text boxes is rich in detail but poorly arranged for a machine to understand. Credit: logoboom/Shutterstock

Medical records are a rich source of health data. When combined, the information they contain can help researchers [better understand diseases](#) and treat them more effectively. This includes COVID-19. But to unlock this rich resource, researchers first need to read it.

We may have moved on from the days of handwritten medical notes, but the information recorded in modern electronic health records can be just as hard to access and interpret. It's an old joke that doctors' handwriting is illegible, but it turns out their typing isn't much better.

The sheer volume of information contained in health records is staggering. Every day, healthcare staff in a typical NHS hospital generate so much text it would take a human an age just to scroll through it, let alone read it. Using computers to analyze all this data is an obvious solution, but far from simple. What makes perfect sense to a

human can be highly difficult for a computer to understand.

[Our team](#) is using a form artificial intelligence to bridge this gap. By teaching computers how to comprehend human doctors' notes, we're hoping they'll uncover insights on how to fight COVID-19 by finding patterns across many thousands of patients' records.

## Why health records are hard going

A significant proportion of a health [record](#) is made up of free text, typed in narrative form like an email. This includes the patient's symptoms, the history of their illness, and notes about pre-existing conditions and medications they're taking. There may also be [relevant information](#) about family members and lifestyle mixed in too. And because this text has been entered by busy doctors, there will also be abbreviations, inaccuracies and typos.

This kind of information is known as unstructured data. For example, a patient's record might say: "Mrs Smith is a 65-year-old woman with atrial fibrillation and had a CVA in March. She had a past history of a #NOF and OA. Family history of breast cancer. She has been prescribed apixaban. No history of hemorrhage."

This highly compact paragraph contains a large amount of data about Mrs Smith. Another human reading the notes would know what information is important and be able to extract it in seconds, but a computer would find the task extremely difficult.

## Teaching machines to read

To solve this problem, we're using something called [natural language processing](#) (NLP). Based on [machine learning](#) and AI technology, NLP algorithms translate the language used in free text into a standardized, structured set of medical terms that can be analyzed by a computer.

These algorithms are extremely complex. They need to understand context, long strings of words and medical concepts, distinguish current events from historic ones, identify family relationships and more. We teach them to do this by feeding them existing written information so they can learn the structure and meaning of language—in this case, publicly available English text from the internet—and then use real medical records for further improvement and testing.

Using NLP algorithms to analyze and extract data from health records has huge potential to change healthcare. Much of what's captured in narrative text in a patient's notes is normally never seen again. This could be important information such as the early warning signs of serious diseases like cancer or stroke. Being able to automatically analyze and flag important issues could help deliver better care and avoid delays in diagnosis and treatment.

### Finding ways to fight COVID-19

By drawing together health records using these tools, we're now using these techniques to see patterns that are relevant to the pandemic. For example, we recently used our tools to discover whether drugs commonly prescribed to treat [high blood pressure](#), diabetes and other conditions—known as [angiotensin-converting enzyme inhibitors \(ACEIs\)](#) and [angiotensin receptor blockers \(ARBs\)](#) – increase the chances of becoming severely ill with COVID-19.

The virus that causes COVID-19 infects cells by binding to a molecule on the cell surface called ACE2. Both ACEIs and ARBs are thought to [increase the amount of ACE2](#) on the surface of cells, leading to concerns that these drugs could be putting people at increased risk from the virus.

However, the information needed to answer this question—how many severely ill COVID-19 patients are being prescribed these drugs—can be recorded both as structured prescriptions and in free text in their [medical records](#). That free text needs to be in a computer-searchable format for a machine to answer the question.

Using our NLP tools, we were able to analyze the anonymised records of 1,200 COVID-19 patients, comparing clinical outcomes with whether or not patients were taking these drugs. Reassuringly, we found that people prescribed ACEIs or ARBs were [no more likely](#) to be severely ill than those not taking the drugs.

We're now expanding how we use these tools to find out more about who is most at risk from COVID-19. For instance, we've used them to [investigate the links](#) between ethnicity, pre-existing health conditions and COVID-19. This has revealed several striking things: that being black or of mixed ethnicity makes you more likely to be admitted to hospital with the disease, and that Asian patients, when in hospital, are at greater risk of being admitted to intensive care or dying from COVID-19.

We've also used these tools to evaluate the [early warning scores](#) that predict which patients admitted to hospital are most likely to become severely ill, and to suggest what additional measures could be used to improve these scores. We're also using the technology to [predict upcoming surges](#) of COVID-19 cases, based on patients' symptoms that doctors have recorded.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

APA citation: Teaching computers to read health records is helping fight COVID-19 (2020, October 19) retrieved 21 September 2022 from <https://medicalxpress.com/news/2020-10-health-covid-.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.*