07
20
COVID-19 EVIDENCE MAP
NEWSLETTER

NIPH
Norwegian Institute of Public Health

From journal paper to map inclusion: The long and winding road for COVID-19 research

The 3405 articles in the Live map of COVID-19 evidence have been through a complicated process. After searching and screening for publications with primary data, each study is carefully categorized by two researchers working independently. Disagreements are reconciled before inclusion in the map.

After a period of internal literature searches, supplemented by material from the Centers for Disease Control and Prevention (CDC), we have just switched to collecting screened results for further processing from the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI Centre) in London. The EPPI Centre is part of the Social Science Research Unit at the Institute of Education, University of London. It is a specialist centre for developing methods for systematic reviewing and synthesis of research evidence. The EPPI Centre searches MEDLINE and Embase and include publications with primary empirical data, systematic reviews or models on COVID-19. We supplement this with searches in SCOPUS.

How is the categorisation organized?

– Next stage in the process is to assign papers to our coders, says Jan Himmels, a member of the research team behind the map. We allocate each set of papers manually to two researchers. The researchers then have access to these studies through their personal EPPI Reviewer account. Once a paper has been categorised, and successfully agreed upon, our EPPI administrator transfers the studies to be included in the next map-update.

– Studies are categorised into 19 publication types and 42 population groups, Jan explains. With the latest update we have 92 topic codes down from the previous 138; this reduction aims to increase efficiency and decrease overlap between separate codes. Even with the reduction in the number of codes, our coding is still very detailed.

The categorisation process is labour-intensive. Depending on the study, it may take 3-15 minutes to code a paper. We have created a coding manual describing all codes in detail, with the aim of reducing ambiguity. Yet, our inclusive approach to coding faces challenges, the breadth of studies we are looking at, brings with it many publications types and topics that are not black or white. Also worth mentioning, is that our coding manual has developed throughout the project to address the developing research. This dynamic approach is great but at times this can require considerable work to realign older codes.

Jan Himmels

Who are the coders?

– We have 28 coders in total, 15 are external volunteers, and 13 are coders from the Norwegian Institute of Public Health, all with a research/ medical background. We have a programme to help coders get on-board with the coding process, with an introduction to the necessary software and our coding manual. Our new coders are then paired up with an experienced coder to continue training on-the-job, with weekly discussions for general questions or specific studies. During these discussion rounds any disagreements in coding are discussed and reconciled. Currently coders are managing to code up to 30 studies in a workday; with our revision of the coding system we hope to increase this number.

What is reconciliation?

– To maintain a high-quality standard two researchers independently categorise all studies. Once both researchers have coded their studies, they discuss their coding and reconcile their differences by agreeing on a code. We have managed to do all coding work through virtual communication. This is both rewarding as well as challenging, as we have volunteers from across the world with their own schedules and time zones. The benefit of this process is a continuous overview over the consistency in our coding efforts. Differences in coding may be due to different viewpoints or human error, but can mostly be resolved without the need of a third researcher to adjudicate.

Machine learning in screening and coding

Use of artificial intelligence (AI) is needed to keep up with the increasing number of COVID-19 publications. Combined with highly skilled human researchers it makes the coding process faster and better – and feasible.

Machine learning is a type of artificial intelligence in which a program "learns" to identify patterns in data, like a scientific paper, with minimal human intervention.

– Minimal human intervention doesn't mean that machine learning is set loose and replaces researchers' work, says Ley Muller, a member of the core team behind the map. It means that we don't have to tell the machine how to do something – it learns how it should do what we ask of it, based on analysing what we have already done manually.

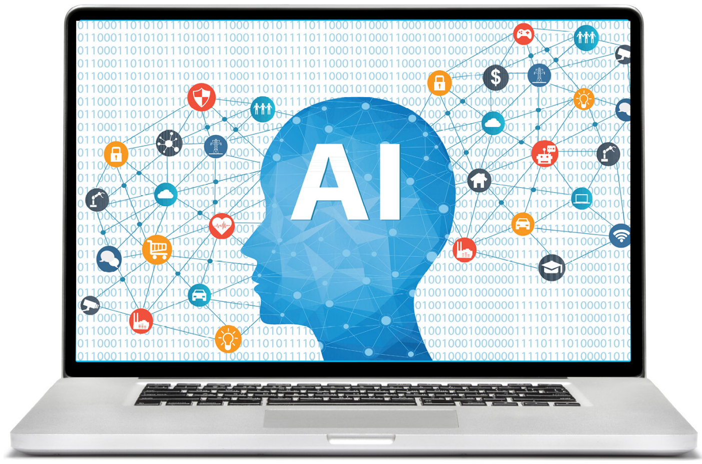
Why is this needed?

– Machine learning is an absolute necessity due to the exponential increase in studies, the demand from decision-makers for as fast as possible evidence synthesis, and the quality control that this technique offers, says Ley.

Researchers make random errors when coding, such as when we get tired or distracted, when we skip over a word, or when we simply misread a sentence. We also make systematic errors, such as when one coder misunderstands a subtopic and therefore codes it incorrectly each time. When we use machine learning, the machine doesn't get tired, distracted, or make other random errors – it only makes the same systematic errors that we do, and once we fix them, we can improve the model!

How does it work in practice?

– In April 2020, we began using machine learning to help us screen incoming studies, says Ley. In the software we use, it meant that machine learning listed high-priority studies first, and studies to exclude last. Simply seeing studies in this logical order allowed us to increase productivity by up to a factor of ten.



We are now also using machine learning to help us code. After we coded 1200 studies, we began testing models against this "gold standard". This helps us to more easily identify case studies from larger observational studies, studies with controls from studies without controls, and so on. We are currently testing a system in which machine learning acts as one coder, and our experienced coders are the second coders, rather than having two human coders. Doing this will cut coding time by half. The more studies we code, the more data we can feed into machine learning, and the better the models become. It's incredibly exciting to see how machine learning can even pick up on some things we don't, such as a study that uses a control group but doesn't use that term in the abstract.

What's nice about machine learning is that all the work we have put in so far by using two human coders goes directly into building better models. It's a good example of machine-human cooperation.

>> Live map of COVID-19 evidence

>> Subscribe to newsletter

RECENT RAPID REVIEWS

1. Should healthcare personnel in nursing homes without respiratory symptoms wear facemasks for primary prevention of COVID-19? A rapid review

Project leader: Gunn Vist
Editor: Magne Nylenna
Date: 10 June 2020
E-mail: Covid-19.evidencemap@fhi.no Co

NIPH
Norwegian Institute of Public Health

helsebiblioteket.no
Norwegian Electronic Health Library