#opendata4covid19

Korea

Seng Chan You, Ajou University

Most slides were created by Dr Yeunsook Rho, HIRA
COVID-19 Pandemic in Korea

Number of confirmed cases (as of 12 March 2020)

- New cases
- Cumulative cases

Cumulative cases: 7869

Days: 1 to 12 March 2020
Epic battle against coronavirus misinformation

**A FACT-CHECKING FRENZY**

Fact-checkers have worked overtime correcting COVID-19 falsehoods. One alliance has collated more than 6,000 examples of fact-checks across a broad range of categories since 14 January. Data as of 19 May.

---

*Stories about what governments and health authorities say, and what is alleged about them.

*Any category that involves a conspiracy theory.
Prof. Chambers (Chair of Center for Open Science and Member of the UK Reproducibility Network Steering Group) said:

“The failure to resolve such basic concerns about the data during the course of normal peer review raises serious questions about the standard of editing at the Lancet and NEJM. If these journals take issues of reproducibility and scientific integrity as seriously as they claim, then they should forthwith submit themselves and their internal review processes to an independent inquiry.”
Open Science

• Open science promotes a more accurate verification of scientific results. By combining the tools of science and information technologies, scientific enquiry and discovery can be sped up for the benefit of society
• Open science reduces duplication in collecting, creating, transferring and re-using scientific material
• Open science results in great innovation potential and increased consumer choice from public research
• Open science promotes citizens’ trust in science. Great citizen engagement leads to active participation in scientific experiments and data collection

https://www.oecd.org/sti/inno/open-science.htm
A new initiative for the collaborative research against COVID-19

OHDSI virtual study-a-thon to support COVID-19 response, to take place 26-29Mar2020...Collaborators wanted!

Team:

In light of the current uncertainty around COVID-19, we have decided to cancel the in-person OHDSI EU Symposium, which was scheduled to take place 27-29Mar2020 in Oxford, UK. An announcement about this is available here. In lieu of this large meeting event, we have decided to coordinate a virtual OHDSI study-a-thon to take place over the same period that will be focused on generating real-world evidence that can inform the current COVID-19 pandemic response.

We expect there are many ways the OHDSI community can contribute to the current situation:

- Characterisation of symptoms and complications of viral diseases
- Prediction of adverse outcomes amongst patients with virus-related hospitalization
- Comparative safety of treatments being considered/used for potential use in COVID-19
#opendata4covid19

Background

- Rapid spread of COVID-19 around the globe, declaration of pandemic by WHO
- Insufficient real-world data and evidence for government to implement policies
- South Korea decided to share the world's first de-identified COVID-19 nationwide patient data with domestic and international researchers.
OHDSI COVID-19
International Study-A-Thon

The meeting will begin shortly.

Collaborating to design and execute observational research and generate real-world evidence to inform the global pandemic

March 26-29, 2020
#opendata4covid19

Objective and Launch

• **Objective**
  – Korea’s nationwide patient Big Data sharing shall help humanity as a whole overcome the disease and produce evidence for effective policy enforcement.

• **Launch of the project**
  – In the regular briefing **on March 27 (Fri)**, the Central Disaster and Safety Countermeasures Headquarters (CDSCHQ) announced the launch of the Global Research Collaboration Project on COVID-19
  – Website address: https://covid19data.hira.or.kr
#opendata4covid19

Data attribute

• **Scope**
  
  – The list of patients of COVID-19 (using submitted claims data) is connected to their history of healthcare service use for the past 3 years, and the entire dataset is de-identified.
  
  • As of May 15, 2020, the total number of claims is 476,508, and the total number of patients is 234,427 (confirmed cases 7,590 patients).
  
  • Data secured for this project will be updated on a regular basis

  – All data has been provided in two format
  
  • Original HIRA data format
  
  • OMOP CDM version 5.3
#opendata4covid19

Research Process

- Distributed research network model
  - Before transferring results, researchers should submit 1-page research proposal, written data use agreement, IRB approval/waiver
  - Prior to any form of formal release of research outcome, the content must be shared with MoHW and HIRA of Korea. (disclosure shall be free of charge)
#opendata4covid19

Together with international researchers

- Status of application (as of May 8)
  - 380 projects have applied from 32 countries
  - Topics include characteristics of COVID-19, correlation between underlying conditions and fatality rate, and machine learning analysis to predict severity level of patients
  - In the beginning, CDM analysis was the most popular. Recently, demand for SAS analysis is growing among Korean researchers.
  - To date, 116 projects submitted analysis codes among a total of 380 (CDM: 27 projects, SAS: 49 projects, R: 40 projects)

<table>
<thead>
<tr>
<th></th>
<th>Korea</th>
<th>The U.S.</th>
<th>The U.K.</th>
<th>Israel</th>
<th>Canada</th>
<th>Italy</th>
<th>Australia</th>
<th>Else</th>
</tr>
</thead>
<tbody>
<tr>
<td>Projects</td>
<td>154</td>
<td>123</td>
<td>19</td>
<td>13</td>
<td>12</td>
<td>10</td>
<td>5</td>
<td>44</td>
</tr>
</tbody>
</table>
#opendata4covid19

Challenges

• Adoption and application of a new research method where researchers from all around the world can join and carry out collaboration for specific

• Diversifying analysis tools
  – How to apply open-source analysis tools (R, Python, etc.) on HIRA’s closed intranet environment to open and share data
Thank You
for your time