Reflection on COVID-19 Data and Resources

Xihong Lin

Department of Biostatistics, School of Public Health
Department of Statistics, Faculty of Arts and Sciences
Harvard University
Broad Institute of MIT and Harvard

xlin@hsph.Harvard.edu

@XihongLin
JAMA and Nature: Analysis of 32,000 Wuhan COVID19 cases

Original Investigation
April 10, 2020

Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China

An Pan, PhD; Li Liu, MD, PhD; Chaolong Wang, PhD; et al

Author Affiliations  |  Article Information

Reconstruction of the full transmission dynamics of COVID-19 in Wuhan

Xingjie Hao, Shanshan Cheng, Degang Wu, Tangchun Wu, Xihong Lin & Chaolong Wang

Nature (2020)  |  Cite this article

Tongji School of Public Health
Huazhong Science and Technology University

An Pan  |  Chaolong Wang  |  Tangchun Wu
COVID-19 Feature #1: High Transmissibility: 
\( R_t \) decreases from 3.54 to 0.27 after interventions

\( R_t = \text{Effective Reproductive Number} \)
The spirit of open science and collaboration has emerged as a hallmark of the fight against the pandemic

- **Viral sequencing data**
  - Gisaid and Nextstrain
  - Challenges: sampling bias and rapid access

- **Aggregated data on deaths, cases and vaccination**
  - Johns Hopkins Coronavirus Resource Center data
  - COVID Tracking Project, USA Fact
  - Our World in Data (data of many countries)
  - Challenges: Heterogeneity and limited age-sex-race specific aggregated data

- **A large increase of open access pre-prints**

- **These open access data help launch many public resources, e.g., dashboards**
Data are irregular in many developing countries, e.g., Africa.
How We Feel App: COVID-19 Stimulated More Partnerships between Academia and Industry

August 26, 2020

PI: X Lin

- 750K+ users
- 15M+ responses

Real time data individual data collection is challenging using traditional venues, e.g. grants

Feng Zhang, MIT and Broad Inst

Ben Silbermann, CEO, Pinterest
A Few Recommendations

- Need infrastructure support for equitable COVID data collection and harmonization, data dashboards, and analytic tools and resources
  - These real time data and analytic efforts for public services are hard to get grant support and traditional incentives

- Build global alliance for COVID-19 data and resources and communication
  - Bringing together academia, government, industry, media to create frameworks and standards to enable voluntary, responsible, and secure sharing of COVID-19 data.

  - Frameworks to share aggregated data (open access) and de-identified and secured individual level data, including EHR data, for research purpose (restricted), e.g., through an application, e.g., UK biobank, or federated analysis and consortium

  - Effective communication of data by engaging media (e.g. FT and NYT visualization)