Sharing COVID-19 Epidemiology Data

Version: 1.0

DOI: 10.15497/rda00049

Authors: RDA COVID-19 Epidemiology Working Group

Co-chair Liaison: Priyanka Pillai

Moderators: Claire Austin, Rajini Nagrani

Published: 4th November 2020

Abstract: An immediate understanding of the COVID-19 disease epidemiology is crucial to slowing infections, minimizing deaths, and making informed decisions about when, and to what extent, to impose mitigation measures, and when and how to reopen society.

Despite our need for evidence based policies and medical decision making, there is no international standard or coordinated system for collecting, documenting, and disseminating COVID-19 related data and metadata, making their use and reuse for timely epidemiological analysis challenging due to issues with documentation, interoperability, completeness, methodological heterogeneity, and data quality.

There is a pressing need for a coordinated global system encompassing preparedness, early detection, and rapid response to newly emergent threats such as SARS-CoV-2 virus and the COVID-19 disease that it causes.

The intended audience for the epidemiology recommendations and guidelines are government and international agencies, policy and decision makers, epidemiologists and public health experts, disaster preparedness and response experts, funders, data providers, teachers, researchers, clinicians, and other potential users.

Keywords: COVID-19; supporting output; epidemiology; recommendations; guidelines; data sharing

Language: English

Citation and Download: RDA COVID-19-Epidemiology WG (2020): Sharing COVID-19 Epidemiology Data, Version 1.0. DOI: 10.15497/rda00049

Research Data Alliance RDA COVID-19

Epidemiology sub-Work Group
Co-moderators: Claire C Austin PhD, Rajini Nagrani PhD, and Gabriel Turinici PhD

SHARING COVID-19 EPIDEMIOLOGY DATA

SUPPORTING OUTPUT

Draft v1.0 (2020-11-04)

Check for the most recent version of this document at: doi.org/10.15497/rda00049
See, also:
- *the RDA-COVID19-WG Zotero Library at: doi.org/10.15497/rda00051*
- *the overarching RDA-COVID19-WG Guidelines and Recommendations across all WGs at:* https://doi.org/10.15497/rda00052

CONTENTS

CHANGE LOG

VERSION	DATE	CHANGE
0.051	2020-06-08	Fixed version number inconsistencies
0.051	2020-06-08	Minor edits
0.051	2020-06-08	Updated Figures
0.051	2020-06-08	Updated Annexes 1-3, and 5-6
0.051	2020-06-08	Updated the Bibliography
0.052	**2020-06-15**	Updated Data Sharing in Epidemiology (Focus, Description, Scope, Recommendations, Guidelines, and Table 1)
0.052	2020-06-15	Updated Annex 1 (COVID-19 data sources)
0.052	2020-06-15	Updated the Bibliography
0.053	**2020-06-28**	Updated the Bibliography
0.053	2020-06-28	Added link to RDA-COVID19-WG Zotero Library
0.053	2020-06-28	Added, *"Check for the most recent version of the present document at: doi.org/10.15497/rda00049"*
0.053	2020-06-28	Added recommendation for zoonoses in Section 3.3 (Policy Recommendations)
0.053	2020-06-28	Added guidelines for causal loops in Section 4.6 (Policy Recommendations)
0.053	2020-06-28	Renumbered the common data model guideline from 4.6 to 4.7
0.053	2020-06-28	Updated Annex 1 (COVID-19 data sources)
0.053	2020-06-28	Updated Annex 2 (Questionnaires)
0.053	2020-06-28	Added zoonoses to Annex 4 (Epi data model)
0.053	2020-06-28	Split the Epi-Trac paper (Annex 5) into two papers (Epi-TRAC and Causal Loops)
0.053	2020-06-28	Updated Annex 5 (Epi-Trac).
0.053	2020-06-28	Updated Annex 6 (Causal loops).
0.053	2020-06-28	Updated Annex 7 (CDM and Epi-STACK).
0.053	2020-06-28	Miscellaneous minor edits throughout
0.053b	**2020-06-30**	Updated link to the new DOI for the final edition (version 1) of the RDA-COVID19-WG Recommendations and guidelines published on 2020-06-30. This is the DOI for the RDA overarching document, not the present supporting output which has not yet been finalized.
0.053c	**2020-06-30**	Fixed DOI error and other miscellaneous small errors.
0.06a	2020-08-31	Updated Annexes 1, 2, 3, 4, 5, 6, and 7.
0.06a	2020-08-31	Annex #2 (Questionnaires) now available on SSRN preprint server.
0.06a	2020-08-31	Annex #3 (Privacy) now available on SSRN preprint server.
0.06a	2020-08-31	Annex #3: Corrected hyperlink to full text.
0.06b	**2020-09-13**	Annex #1 (COVID-19 data sources): Updated supplementary tables S1-S8.
0.06b	2020-09-13	Annex #6 (Causal loops) now available on SSRN preprint server.
0.06b	2020-09-13	Miscellaneous minor edits.
0.06c	**2020-09-21**	Miscellaneous minor edits.
0.06c	2020-09-21	Annex #1 (Data sources): Updated text and tables
0.06c	2020-09-21	Annex #1 (Data sources) now available on SSRN preprint server.
1.0	**2020-11-03**	**Final version**

DATA SHARING IN EPIDEMIOLOGY

1 Focus and Description

An immediate understanding of the COVID-19 disease epidemiology is crucial to slowing infections, minimizing deaths, and making informed decisions about when, and to what extent, to impose mitigation measures, and when and how to reopen society.

Despite our need for evidence based policies and medical decision making, there is no international standard or coordinated system for collecting, documenting, and disseminating COVID-19 related data and metadata, making their use and reuse for timely epidemiological analysis challenging due to issues with documentation, interoperability, completeness, methodological heterogeneity, and data quality.

The intended audience for the epidemiology recommendations and guidelines are government and international agencies, policy and decision makers, epidemiologists and public health experts, disaster preparedness and response experts, funders, data providers, teachers, researchers, clinicians, and other potential users.

2 Scope

Epidemiology underpins COVID-19 response strategies and public health measures. The recommendations and guidelines support development of an internationally harmonized specification to enable rapid reporting and integration of epidemiology and related data across domains and between jurisdictions

The guidelines outline a data driven, coordinated global system that encompasses preparedness, early detection, and rapid response to newly emergent threats such as SARS-CoV-2 virus and the COVID-19 disease that it causes.

Supporting Output

The present supporting output provides supplemental resources, and further develops the global data-driven vision described in the guidelines. This includes a proposed computable framework to support system responses for emerging pathogens. It offers compatible and reliable data models, protocols, and action plans for newly identified threats such as COVID-19.

3 Policy Recommendations

3.1 General

1. Urgently update data sharing policies and Memoranda of Understanding (MOUs) across all domains, in government, healthcare systems, and research institutions to support Open Data, Open Science, scientific data modernization, and linked data life cycles that will enable rapid and credible scientific and epidemiologic discovery, and fast-track decision-making.
2. Streamline data flow between sub-national jurisdictions/institutions and their national government, and countries and international organizations.

3. Implement a "*data first*" publication policy in research by treating publication of data articles in "open" peer-reviewed data journals, including the deposit of data and associated code in a trusted digital repository with tiered access to appropriately credentialed people and machines to preserve data security.
4. Peer-reviewed data articles should be treated as first-class research outputs equal in value to traditional peer-reviewed articles.
5. Call upon the international Open Government Partnership (OGP) to add "*Open Science*" as one of its Policy Areas to be included in National Action Plans. Hold member countries accountable for developing and implementing Open Science commitments.
6. Publish situational data, analytical models, scientific findings, and reports used in decision-making and justification of decisions (OGP 2020).

3.2 Information Technology and Data Management

Properly funded state of the art infrastructure is required to support advanced research, as well as the and data management and data sharing required for rapid response and collaboration (See Infrastructure Investment). For epidemiology in particular:

1. Ensure an appropriate semantic annotation of data to facilitate its comparability across studies and countries, using as much as possible established standards (e.g. LOINC, UMLS).
2. Rapidly develop standardised tools for aggregating microdata to a harmonised format(s) that can be shared and used while minimising the re-identification risk for individual records.
3. Develop machine readable citations and micro-citations for dynamic data. Rapid development of: (a) Resolvable Persistent Identifiers, rather than Uniform Resource Locators (URLs); (b) Machine readable citations; (c) Micro-citations that refer to the specific data used from large datasets; and, (d) Date and Time Access citations for dynamic data (ESIP, 2019).

3.3 COVID-19 Epidemiological data, analysis, and modeling

1. Implement a global system for early detection and rapid response to emerging zoonoses, integrated across systems for reporting human and animal diseases as well as their vectors and geographical distribution ((CDC 2020; CGH 2019; eCDC 2019; GLEWS 2006; NASEM 2008, 2009a,b; WHO 2017; WHO, FAO and OIE 2019).
2. Catalogue and document all zoonotic diseases with associated reservoir species and vectors to establish and maintain a global database with potential risks related to humans.
3. Rapidly develop a consensus standard for COVID-19 surveillance data:
    a. Definition of and reporting criteria for COVID-19 testing, reporting on testing, and testing turnaround times.
    b. Policies and definitions: interventions, contact tracing, reporting of cases, deaths, hospitalisations and length of stay, ICU admissions, recoveries, reinfections, time from contact if known, symptoms onset and detection, through clinical course and interventions, to death or recovery, comorbidities, long-term effects in recovered cases, sequelae and immunity, location, demographic, socioeconomic information, and outcome of resolved cases.
    c. Uniform standard daily reporting cut-off time.

4. Rapidly develop an internationally harmonised specification to enable the export/import/integration of epidemiologic data across different levels of data generation (e.g., clinical systems, population-based surveillance/research data, data from biomarker and omics studies, death certification, health insurance data), and successful record-linkage.
5. Develop systems that support workflows to link and share data between different domains, while protecting privacy and security. Use domain specific, time stamped, encrypted person identifiers for this purpose based on industry-standard encryption and cryptographic constructions.
6. Implement internationally harmonised COVID-19 intervention protocols based on peer-reviewed empirical modelling and epidemiological evidence, considering local conditions.
7. Publish situational data, analytical models, scientific findings, and reports used in decision-making and justification of decisions (OGP, 2020b).
8. Account for public health decision making demands in COVID-19 studies.
9. Harmonise approaches to comparably assess and quantify side-effects of pandemic containment and mitigation measures.
10. Report underlying assumptions and quantify effects of uncertainties on all reported parameters and conclusions for all model predictions etc.
11. Implement a data-driven approach for early identification of hotspots.

4 Guidelines

These guidelines highlight current system challenges and offer solutions to help support a larger framework designed to coordinate and structure the collection and use of COVID-19 related data. Six focus areas, described in the guidelines and supporting output (data sources, instruments, privacy, epidemiological data model, computable framework, and an epi-stack framework), progressively develop a data driven global vision for managing novel biological threats such as COVID-19. We begin with population level data sources that drive the public health strategy and response at all stages of the COVID-19 threat, from emergence through containment, mitigation, and reopening of society. We then survey clinical and population-based instruments that collect data and discuss preservation of individuals' privacy in shared COVID-19 related data. A full spectrum data model is presented encompassing hospital specific surveillance and electronic health records together with field-based demographic and epidemiological surveillance. We propose Epi-TRACS, a computable framework for emerging pathogen action plans, and an epi-stack that uses the Common Data Model with COVID-19 to integrate clinical data and epidemiological data.

4.1 COVID-19 Population Level Data Sources

Although jurisdictions within countries send COVID-19 population level data to the national level, and member countries send data to the WHO, other organisations also collect COVID-19 surveillance data from various sources for a variety of reasons (Table 1). Epidemiologists are thus faced with a situation where it is difficult to assess which datasets are the most up-to-date, complete, and reliable.

See ANNEX 1 for further details and discussion.

Table 1. COVID-19 population level data sources

SOURCE	DATA
Allen Institute for AI	COVID-19 Open Research Dataset (CORD-19)
Apple Inc	COVID-19-Mobility Trends Reports
European Centre for Disease Control	Geographic distribution of COVID-19 cases worldwide
European Centre for Disease Control	The European Surveillance System (TESSy).
Institute for Health Metrics and Evaluation (IHME)	Global Health Data Exchange (GHDx)
Google Inc.	COVID-19 Community Mobility Report
Johns Hopkins University	COVID19 dataset
Kieren Healy Rpackage	Rpackage - COVID19 Case and Mortality Time Series
University of Oxford	COVID19 dataset
The Atlantic	Tracking Project
The New York Times	Covid-19 Data in the United States
U.N.	Humanitarian Data Exchange (HDX)
U.S. Centre for Disease Control	Cases of COVID19 in the U.S.
University of Washington	Be Outbreak Prepared
World Bank	Understanding the Coronavirus (COVID-19) pandemic through data
World Health Organization (WHO)	Novel Coronavirus (2019-nCoV) situation reports
Worldometer	COVID19 data

4.2 Interoperable COVID-19 epidemiological surveillance: Clinical and Population-based instruments

International efforts are currently underway to create COVID-19 instruments/questionnaires (Tables 2 and 3). These COVID-specific tools are concentrated at person-level for clinic/hospital surveillance (e.g., Case Report Forms-CRFs), or community surveillance (e.g., questionnaire for general population), and do not necessarily collect the same data. Adherence of new studies to already introduced instruments will strongly enhance the comparability of results.

Table 2. Questionnaire instruments: Reference studies

COUNTRY	QUESTIONNAIRE
CLINICAL	
Australia	NSW Case questionnaire
Austria	EMS
Europe	TESSy
Germany	Covid-19 research dataset
Uganda	Perinatal COVID-19 Uganda
US	Human Infection with 2019 Novel CoronavirusPerson Under Investigation (PUI) and Case Report Form
Worldwide (WHO member states)	Global COVID-19: clinical platform: novel coronavius (COVID-19): rapid version
POPULATION BASED	
Brazil	Brazil Prevalence of Infection Survey
Europe	Questionnaire by WHO Europe
Germany	GESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in Germany
Germany	NAKO COVID-19 Survey tool

Israel	One-minute population wide survey
LMICs	LMIC Covid Questionnaire
South Africa	South African Population Research Infrastructure (SAPRIN) COVID-19 Screening Form
South Asian Countries	National Institute for Health Research (NIHR) Global Health Research Unit
UK	UK COVID-19 Questionnaire
Worldwide (WHO)	Population-based age-stratified sero-epidemiological investigation protocol for COVID-19 virus infection

Table 3. Questionnaire instruments: Resources

INITIATIVE
NIH Public Health Emergency and Disaster Research Response (DR2)
COVID-19OBSSR Research Tools
PhenX COVID-19 Toolkit

Some of the questionnaire initiatives shown in Tables 2 and 3 are currently feeding into the construction of a COVID-19 demographic and epidemiological surveillance question bank that can be used to form locality specific surveys with both common and distinct questions by domains and cohorts (Wellcome Trust). Some, such as the UK COVID-19 Questionnaire, or the Covid-19 research dataset are now being funded. Question banks, once they become operational can be queried and filtered by domain, cohort, question text, etc. Based on such queries, new questionnaire products can be developed that are more or less interoperable, depending on the questions selected and the capture of "localisation" information in the question metadata when questions are reused from one survey to the next.

See ANNEX 2 for further details and discussion.

4.3 Preservation of individuals' privacy in shared COVID-19 related data

Data sharing is essential to improve epidemiological analysis, cross-border pandemic modelling, and coordinated policy development between countries. To ensure privacy, both pseudo-anonymization of direct identifiers (e.g. patient specific ID's) and anonymization of indirect identifiers (e.g. socio-demographic information on individuals) must be applied. In addition, it is necessary to control statistical disclosure risk to prevent identification of individuals and their health status using a combination of indirect identifiers such as education level, sex, age, and clinical condition, among others (Duncan et al., 2011; Templ et al., 2015; Templ, 2017). Using synthetic data may be an option to lower re-identification risks while retaining properties of the original data sets.

See ANNEX 3 for further details and discussion.

4.4 A Full Spectrum View of the COVID-19 Data Domain: An Epidemiological Data Framework

The COVID-19 epidemiology that guides public health decisions is dependent on interoperable input data from across a wide variety of domains that include not only clinical, surveillance, research, and modelling data, but also administrative, demographic, socioeconomic, cultural practices and lifestyle, and environmental data, amongst others.

An epidemiological surveillance data model must include the primary data domains that need to be integrated to understand COVID-19, and to improve surveillance and follow-up: (a) clinical event history and disease milestones; (b) epidemiological indicators and reporting data; (c) contact tracing; (d) personal risk factors.

Standardization challenges within each of these domains remain to be solved before data can be effectively integrated across domains for epidemiology studies. For example, on the clinical side, the U.S. Clinical Data Interchange Standards Consortium (CDISC) new specification (Interim User Guide for COVID-19), and the WHO Core and Rapid COVID-19 Case Reporting Forms used in low- and middle-income Countries (LMIC) require additional harmonization.

See ANNEX 4 for further details and discussion.

4.5 Epi-TRACS: A data framework for rapid detection and whole system response for emerging pathogens

WHO's Global Influenza Surveillance Response System (GISRS) is a well-established network of more than 150 national public health laboratories in 125 countries that monitors the epidemiology and virologic evolution of influenza disease and viruses (WHO, 2020).

Prior to the COVID-19 outbreak, WHO was already engaged in re-examining GISRS's long-term fitness-for-purpose. In line with these short-term considerations, and with GISRS long-term aspirations, we are recommending a real time, adaptable, rapid response system that supports developing countries, and that employs new technology to combat pandemics and other emerging diseases. The RDA-COVID19-Epidemiology group recommends the creation of a WHO-led EPIdemiological Translational Research Action Coordination System (Epi-TRACS) to add an implementation layer to the existing WHO policies, guidelines, partnerships, and information exchange stack adapted to country-specific contexts.

See ANNEX 5 for further details and discussion.

4.6 COVID-19 Emergency public health and economic measures causal loops: A computable framework

Causal loop modeling may be valuable in assessing system sustainability and system resiliency (Bahri 2020; Ricciardi et al. 2020; Wicher 2020). A computable framework in which the actions taken in response to COVID-19 sentinel surveillance can be simulated and assessed both retrospectively and prospectively based on a causal loop diagram may help inform decision making.

See ANNEX 6 for further details and discussion.

4.7 Common Data Models and Full Spectrum Epidemiology: An Epi-STACK framework for COVID-19 epidemiology datasets

Common Data Model (CDM)

Data models may make use of the broad ecosystem of surveillance and clinical data that can also include contact tracing apps, biospecimens, and environmental sample data collected in the community/population or clinic/hospitals.

An emulated trials approach may enable assessment of various risk and prognostic factors (Hernán et.al., 2016). Application of a Common Data Model (CDM) for COVID-19 would facilitate comparing clinical burden and patient outcomes in the context of previous environmental and exposures and comorbidities.

Another possible use case is decision support following an early warning system alert of emergence of a novel pathogen such as SARS-CoV-2. The CDM provides a framework for making public health policy decisions, using partial information about the pandemic that leverages population-level population and health information, person-level epidemiological surveillance information collected in the field and, at the same time or alternatively, person-level patient care information collected in a clinic or hospital setting.

Epi-Stack

The WHO has established the Information Network for Epidemics (Epi-WIN) covering four strategic areas: (a) Identify; (b) Simplify; (c) Amplify; and, (d) Quantify. Evidence is gathered, appraised, and assessed to help form recommendations and policies that have an impact on the health of individuals and population.

The RDA Epi subWG proposes an expanded Epi-Stack feeding into Epi-WIN. This would bring together in a managed system a common data model, the epidemiological surveillance data model, clinical and questionnaire data, population level indicators, and core use cases (Epi-TRACS early warning and response system, decision support research, and patient care research). It is critical that resiliency be built into the system, from the standpoint of how the system functions under stress. When the degree of complexity and interdependencies increase in human made systems, there is always the risk of collapse if not enough balance is built into the system (both the IT infrastructure, data governance, and the "people" part).

See ANNEX 7 for further details and discussion.

Acknowledgments

The RDA-COVID19-WG would like to acknowledge the following people who have contributed their time, knowledge, and expertise to generate these recommendations and guidelines. Listed alphabetically by last name and ORCID ID:

Claire C. Austin (0000-0001-9138-5986); Marlon L Bayot (0000-0002-5328-150X); Maeve Campman (0000-0002-6613-7144); David Delmail (0000-0003-2836-6496); Cyrille Delpierre (0000-0002-0831-080X); Gayo Diallo (0000-0002-9799-9484); Jay Greenfield (0000-0003-2773-5317); Chifundo Kanjala (0000-0003-0540-8374); Birte Lindstaedt (0000-0002-8251-1597); Gary Mazzaferro (0000-0002-3196-7201); Robin Michelet (0000-0002-5485-607X); Daniel Mietchen (0000-0001-9488-1870); Rajini Nagrani (0000-0002-1708-2319); Carlos Luis Parra-Calderón (0000-0003-2609-575X); Priyanka Pillai (0000-0002-3768-8895); Stefan Sauermann (0000-0003-0824-9989); Carsten Oliver Schmidt (0000-0001-5266-9396); Meg Sears (0000-0002-6987-1694); Henri Tonnang (0000-0002-9424-9186); Gabriel Turinici (0000-0003-2713-006X); and, Anna Widyastuti (0000-0003-2149-935X).

ANNEX 1 – **COVID-19 Surveillance data and models: Review and analysis**

Claire C. Austin[1,§], Anna Widyastuti[2]; Nada El Jundi[3], Rajini Nagrani[3,§]; and the RDA-COVID19-WG[4]

[1]Environment and Climate Change Canada, [2]Hasselt University, Belgium, [3]Leibniz Institute for Prevention Research and Epidemiology-BIPS, Germany, [4]This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

All views and opinions expressed are those of the co-authors, and do not necessarily reflect the official policy or position of their respective employers, or of any government, agency or organization.

[§]Corresponding authors: claire.austin@canada.ca and nagrani@leibniz-bips.de

CITE AS: Claire C. Austin, Anna Widyastuti, Nada El Jundi, Rajini Nagrani; and the RDA-COVID19-WG (2020). *COVID-19 Population level data sources: Review and analysis.* 37 pages. Version 1.0. http://dx.doi.org/10.2139/ssrn.3695335

ABSTRACT

Background. Reliable COVID-19 data are critical for understanding the disease and spread of the pandemic, for decision-making, for developing and implementing public health measures, and for tracking the effectiveness of interventions. Currently, however, there is a confusing plethora of publicly available COVID-19 surveillance data resources. Relevant websites are frequently poorly designed making it extraordinarily time-consuming and frustrating to find and extract the relevant information.
Methods. A systemic search of government, official agency, and non-government sources of COVID-19 surveillance and related data, computer code, and forecasting models was conducted.
Results. A comprehensive compendium was built of COVID-19 surveillance data and models having worldwide national coverage, and some sources of particular interest having sub-national coverage. Hyperlinks are provided to download data or computer code from each of the resources. For each resource, a concise description of the data and metadata, including identification of the data sources used to compile the data is provided. The compendium is provided in the supplementary material, organized in nine tables: (1) COVID-19 surveillance datasets and sources; (2) Databases or catalogues of COVID-19 surveillance data; (3) Resources that provide a corpus of COVID-19 related text; (4) Resources that track COVID-19 government responses; (5) R code potentially useful for analysis of COVID-19 data; (6) COVID-19 related data analysis platforms; (7) COVID-19 models; (8) Useful visualizations of COVID-19 data that go beyond the usual '*dashboards*'; and, (9) Commercial sites that showcase their product with a COVID-19 use case. Selected examples of data resources and models are provided in two additional tables in the body of the text.
Conclusion. There is no single source of truth for COVID-19 surveillance data. Government and non-government data were found to be fragmented and difficult to find, access, and interoperate. There is an urgent need to develop a common standard for reporting communicable disease surveillance data based on FAIRER (Findable, Accessible, Interoperable, Reusable, Ethical, and Reproducible) data.

KEYWORDS: COVID-19, data, surveillance, model, sources, computer code, R, FAIRER data.

PLEASE SEE THE SSRN PREPRINT FOR THE FULL TEXT (37 pages)

ANNEX 2 –**COVID-19 Questionnaires, surveys, and item banks: Overview of clinical- and population-based instruments**

Carsten O. Schmidt[1,*, §] Rajini Nagrani[2,*] Christina Stange[1], Matthias Löbe[3], Atinkut Zeleke[1], Guillaume Fabre[4], Sofiya Koleva[3], Karine Trudeau[4], Stefan Sauermann[5], Jay Greenfield[6], Claire C Austin[7]; and the RDA-COVID19-WG[8]

[1]University Medicine Greifswald, Germany, [2]Leibniz Institute for Prevention Research and Epidemiology-BIPS, Germany, [3]Institute for Medical Informatics, Statistics and Epidemiology, University of Leipzig, Germany, [4]Maelstrom Research Institute, McGill University, Canada, [5]UAS Technikum Wien, Austria, [6]Data Documentation Initiative, USA, [7]Environment and Climate Change Canada; [8]This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG   Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

All views and opinions expressed are those of the co-authors, and do not necessarily reflect the official policy or position of their respective employers, or of any government, agency, or organization.

* Co first authors: Schmidt and Nagrani

§ Corresponding author: carsten.schmidt@uni-greifswald.de

CITE AS: Carsten O. Schmidt*, Rajini Nagrani*,§, Christina Stange, Matthias Löbe, Atinkut Zeleke, Guillaume Fabre, Sofiya Koleva, Stefan Sauermann, Jay Greenfield, Claire C. Austin; and the RDA-COVID19-WG (2020). *COVID-19 Questionnaires, surveys and item-banks: Overview of clinical- and population-based instruments.* 9 pages. Version 1.0 Available at: http://dx.doi.org/10.2139/ssrn.3652027

ABSTRACT

New COVID-19 related instruments are being rapidly developed around the world to collect patient- and population-based information. Heterogeneity between instruments limits comparability of results. The present study provides an overview of instruments and resources. We scoped the content domain on a selection of instruments using the Maelstrom taxonomy. Content of the instruments varied widely, from proximal measures (e.g., clinical symptoms, comorbidities, etc.) to distal measures such as political attitudes. We recommend that researchers reuse existing instruments to the greatest extent possible, and that they make results openly available in machine-readable format to facilitate reuse and maximize comparability of results across studies and countries.

Keywords: COVID-19, SARS-CoV-2, questionnaire, survey, collection tool, instrument, epidemiology, clinical, population, symptoms, interoperability, taxonomy, semantic annotation.

PLEASE SEE THE SSRN PREPRINT FOR THE FULL TEXT (9 pages)

ANNEX 3 – **Preservation of individuals' privacy in shared COVID-19 related data**

Stefan Sauermann[1], Chifundo Kanjala[2], Matthias Templ[3,*], Claire C. Austin[4]; and the RDA-COVID19-WG[4]

[1]UAS Technikum Wien, [2]London School of Hygiene and Tropical Medicine, [3]Zurich University of Applied Sciences, [4]Environment and Climate Change Canada, [5]This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing; and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

All views and opinions expressed are those of the co-authors, and do not necessarily reflect the official policy or position of their respective employers, or of any government, agency or organization.

Corresponding author: chifundo.kanjala@lshtm.ac.uk

CITE AS: Stefan Sauermann, Chifundo Kanjala§, Matthias Templ, Claire C. Austin; and the RDA-COVID19-WG. (2020). *Preservation of individuals' privacy in shared COVID-19 related data.* 13 pages. Version 1.0. Available at: http://dx.doi.org/10.2139/ssrn.3648430

ABSTRACT

This paper provides insight into how restricted data can be incorporated in an open-be-default-by-design digital infrastructure for scientific data. We focus, in particular, on the ethical component of FAIRER (Findable, Accessible, Interoperable, Ethical, and Reproducible) data, and the pseudo-anonymization and anonymization of COVID-19 datasets to protect personally identifiable information (PII). First, we consider the need for the customisation of the existing privacy preservation techniques in the context of rapid production, integration, sharing and analysis of COVID-19 data. Second, the methods for the pseudo-anonymization of direct identification variables are discussed. We also discuss different pseudo-IDs of the same person for multi-domain and multi-organization. Essentially, pseudo-anonymization and its encrypted domain specific IDs are used to successfully match data later, if required and permitted, as well as to restore the true ID (and authenticity) in individual cases of a patient's clarification. Third, we discuss application of statistical disclosure control (SDC) techniques of COVID-19 disease data. To assess and limit the risk of re-identification of individual persons in COVID-19 datasets (that are often enriched with other covariates like age, gender, nationality, etc.) to acceptable levels, the risk of successful re-identification by a combination of attribute values must be assessed and controlled. This is done using statistical disclosure control for anonymization of data. Lastly, we discuss the limitations of the proposed techniques and provide general guidelines on using disclosure risks to decide on appropriate modes for data sharing to preserve the privacy of the individuals in the datasets.

KEYWORDS: Pseudo-anonymization, statistical disclosure control, data anonymization, data sharing, privacy, personally identifiable information, PII, COVID-19, open science.

PLEASE SEE THE SSRN PREPRINT FOR THE FULL TEXT (13 pages)

DISCUSSION PAPER

ANNEX 4 – **Full Spectrum View of the COVID-19 data domain: An Epidemiological Data Framework**

Jay Greenfield[1, §], Rajini Nagrani[2], Meg Sears[3], Gary Mazzaferro[4], Claire C Austin[5, §]; and the RDA-COVID19-WG[6]

[1]Data Documentation Initiative, USA, [2]Leibniz Institute for Prevention Research and Epidemiology-BIPS, [3]Ottawa Hospital Research Institute, [4]Consultant, [5]Environment and Climate Change Canada, [6]This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

[§] Corresponding authors: nightcleaner@gmail.com and claire.austin@canada.ca

CITE AS*:* Jay Greenfield, Rajini Nagrani, Meg Sears, Gary Mazzaferro, Claire C. Austin; and the RDA-COVID19-WG. (2020). *A Full Spectrum View of the COVID-19 data domain: An Epidemiological Data Framework.* [IN: Sharing COVID-19 epidemiology data, Research Data Alliance RDA COVID-19 Epidemiology WG]. Research Data Alliance RDA-COVID19-Epidemiology WG. https://doi.org/10.15497/rda00049

ABSTRACT

We propose an epidemiological surveillance conceptual data framework to guide construction of a Common Data Model and database schema for COVID-19 epidemiological research. The framework encompasses hospital specific surveillance in line with the WHO COVID-19 core and rapid case report forms, electronic health records, and field-based COVID-19 demographic and epidemiological surveillance.

KEYWORDS*: COVID-19, epidemiology, disease milestones, contacts, personal risk factors.*

BACKGROUND

This is the first paper in a series of four discussion papers that progressively build a proposed Epi-STACK framework to support full spectrum COVID-19 epidemiology. Based on the traditional, deterministic Susceptibility, Exposure, Infection, Recovery (SEIR) model, the present paper introduces the concept of a full spectrum epidemiological surveillance framework. The second paper in the series proposes a conceptual early warning and response framework for development of an, "*Epidemiology Translational Research Action Coordinated System* (Epi-TRACS). The third paper proposes the use of systems thinking and causal loop diagrams (CLD) to better understand multiple interrelated factors and impacts. The last paper in the series proposes an Epi-STACK framework for development of COVID-19 datasets based on the Common Data Model to support full spectrum epidemiology to support public health recommendations. This series of four papers follows upon a review and analysis of COVID-19 population level data sources (Austin et al. 2020), an overview of COVID-19 clinical and population based questionnaire and survey instruments (Schmidt et al. 2020), and explores how restricted data can be incorporated in an open-be-default-by-design digital infrastructure for scientific data with preservation of individuals' privacy in the COVID-19 context (Sauermann et al. 2020).

The full spectrum COVID-19 domain framework[1] accounts for the individual's entire experience, including susceptibility, exposure, disease severity, infection, treatment, sequelae, and either death or recovery (Pesquita et al. 2014; WHO 2020 Apr 23; WHO 2020 Mar 24). Susceptibility includes person-level risk factors, public health measures, and individuals' compliance with public health orders and recommendations.

A FULL SPECTRUM EPIDEMIOLOGY FRAMEWORK

We propose development of an augmented full spectrum data framework based on the Susceptibility, Exposure, Infection, Recovery (SEIR) model (Sun and Kang 2020). The framework would take into account public health measures (Giordano et al. 2020) and differences in the availability of resources between High Income Countries (HIC) and Low and Middle Income Countries (LMIC). In HICs, the SEIR model may be more informed by the hospital experience, while in LMICs it may be community-based demographic and epidemiological surveillance that carry greater influence (Wang et al. 2020). Some LMICs also benefit more directly from lessons learned from prior experience with Ebola and HIV AIDS. A COVID-19 epidemiological surveillance data framework would incorporate clinical, community, and epidemiological surveillance data (Figure 1).

Figure 2 breaks out the framework into three components: (A) disease milestones; (B) contacts; and, (C) personal risk factors, and identifies their provenance. Disease milestones and the personal risk factors components would be based on the WHO Global Influenza Surveillance and Response System (GISRS 2020a), COVID-19 sentinel surveillance (WHO 2020b), Wellcome Trust Longitudinal Populations Strategy (Wellcome Trust 2017), WHO COVID-19 core version and rapid version case report forms (CRF) (WHO 2020b; WHO 2020c), and the Wellcome Trust COVID-19 LPS HIC and LMIC Questionnaires (Wellcome Trust COVID-19 LPS Questionnaires).

[1] In ontology engineering, a domain model is a formal representation of a knowledge domain with concepts, roles, data types, individuals, and rules, typically grounded in a description logic.

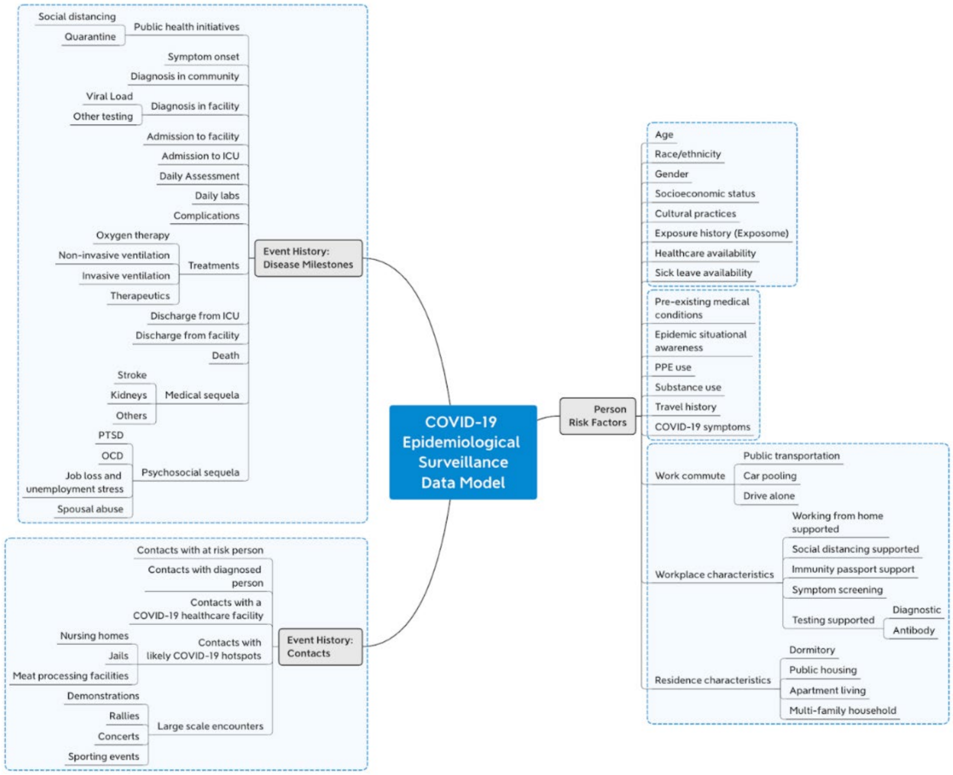

Figure 1: COVID-19 epidemiological surveillance data framework concept

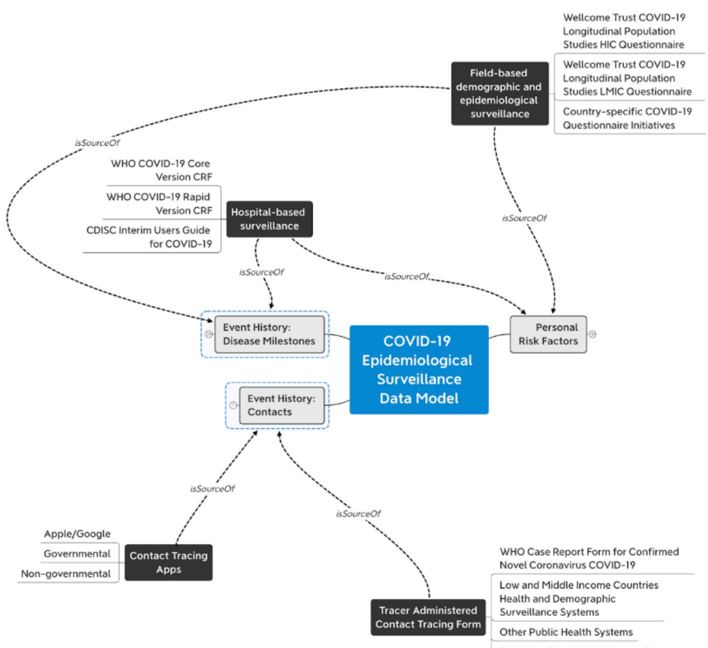

Figure 2: Overview of the proposed COVID-19 epidemiological surveillance data framework

DISCUSSION

The proposed epidemiological surveillance data framework would inform development of an extensible database schema. See, for example, schema.org (2020) and HL7 FHIR (2020). The proposed framework is longitudinal in scope, and the extensible schema would need to be longitudinal as well to account for such things as repeat visits to health care facilities (e.g., hospitals, clinics, and nursing homes), changed diagnoses, multiple diagnostic tests, and repeat encounters with field workers. New contact tracing tools will need to be developed for large scale encounters where traditional methods fail. The personal risk factors component would be refined as greater understanding is gained about COVID-19 disease, how it spreads, and the effectiveness of different public health measures.

Author roles

All authors accept responsibility for the content of the article.
Conceptualization: JG, MS, GM, CCA. Methodology: JG, GM. Investigation: JG, RN, MS, RN, GM, CCA. Validation: RN, MS, GM, CCA. Writing: JG, MS. Visualization: JG

Funding: None.
Conflict of interest: None to declare.

REFERENCES

Claire C. Austin, Anna Widyastuti, Nada El Jundi, Rajini Nagrani; and the RDA-COVID19-WG (2020). *COVID-19 Population level data sources: Review and analysis.* http://dx.doi.org/10.2139/ssrn.3695335

Giordano G, Blanchini F, Bruno R et al. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nat Med. 2020 Jun;26(6):855-860.* https://doi.org/10.1038/s41591-020-0883-7

HL7 (2020). Introduction to HL7 FHIR. https://hl7.org/FHIR/summary.html.

Pesquita C, Ferreira JD, Couto FM, Silva MJ (2020). The epidemiology ontology: an ontology for the semantic annotation of epidemiological resources. J Biomed Semantics. 2014;5(1):4. Published 2014 Jan 17. doi:10.1186/2041-1480-5-4

Schema.org (2020). Organization of Schemas. https://schema.org/docs/schemas.html

Sun, P and Kang L. (2020). An SEIR Model for Assessment of Current COVID-19 Pandemic Situation in the UK. https://doi.org/10.1101/2020.04.12.20062588.

Wang C, Pan R, Wan X, et al. (2020). A longitudinal study on the mental health of general population during the COVID-19 epidemic in China [published online ahead of print, 2020 Apr 13]. *Brain Behav Immun*. S0889-1591(20)30511-0. doi:10.1016/j.bbi.2020.04.028

Wellcome Trust (2017). Wellcome's Longitudinal Population Studies Working Group https://wellcome.ac.uk/sites/default/files/longitudinal-population-studies-strategy_0.pdf.

WHO (2020a). Global Influenza Surveillance and Response System (GISRS). World Health Organizationhttps://www.who.int/influenza/gisrs_laboratory/en/.

WHO (2020b). COVID-19 Core Version CRF. World Health Organization Core COVID-19 CRF.
WHO (2020c). COVID-19 Rapid Version CRF. World Health Organization RAPID COVID-19 CRF

DISCUSSION PAPER

ANNEX 5 – **Epi-TRACS: Rapid detection and whole system response for emerging pathogens such as SARS-CoV-2 virus and the COVID-19 disease that it causes**

Jay Greenfield[1, §], Henri E.Z. Tonnang[2], Gary Mazzaferro[3], Claire C. Austin[4, §]; and the RDA-COVID19-WG[5]

[1]Data Documentation Initiative, USA, [2]International Center of Insect Physiology and Ecology (ICIPE), Kenya, [3]Consultant, USA, [4]Environment and Climate Change Canada, [5] This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

[§] Corresponding authors: nightcleaner@gmail.com and claire.austin@canada.ca

CITE AS: Jay Greenfield, Henri E.Z. Tonnang, Gary Mazzaferro, Claire C. Austin; and the RDA-COVID19-WG. (2020). *Epi-TRACS: Rapid detection and whole system response for emerging pathogens such as SARS-CoV-2 virus and the COVID-19 disease that it causes.* [IN: Sharing COVID-19 epidemiology data, Research Data Alliance RDA COVID-19 Epidemiology WG]. https://doi.org/10.15497/rda00049

ABSTRACT
Despite repeated warnings and recommendations to prepare for emerging pathogens that could result in pandemics, the world was not prepared to respond to the fast moving threat of COVID-19. The present paper proposes a conceptual early warning and response framework for "Epidemiology Translational Research Action Coordinated System" (Epi-TRACS).

KEYWORDS: COVID-19, epidemiology, surveillance, economy, causal analysis

BACKGROUND

COVID-19 threat detection was slow and ineffective, resulting in rapid development of a pandemic. Countries around the world have implemented a disparate series of public health measures in attempting to suppress and mitigate spread of the disease. The world was not prepared to respond to a novel zoonose that spreads with the tempo and severity of COVID-19. The result has been widespread lockdowns which have had serious socioeconomic consequences for both High Income Countries (HIC) and for Low and Middle Income Countries (LMIC) (Gates 2020; Bai et al. 2020).

Previous outbreaks such as H5N1 in 1997, SARS-1 in 2003, H7N2 in 2004, MERS-CoV in 2012, Ebola in 2014, and others, should have been a wake-up call. Failure to heed warnings and recommendations has had expensive and tragic consequences (NASEM 2008, 2009a,b, 2010, 2015, 2018, 2020a,b,c).

Over the past 20 years, before COVID-19, it is estimated that zoonotic diseases cost $100US billion in economic losses, and killed two million people every year - mostly in LMICs (UNEP and ILRI 2020). It has been estimated that the COVID-19 pandemic will cost $9 trillion over the next few years (UNEP and ILRI 2020).

In 2008, the National Academies of Sciences, Engineering, and Medicine (NASEM) conducted a workshop on achieving sustainable global capacity for surveillance and response to emerging diseases of zoonotic origin:

> *"These newly identified diseases have emerged primarily as a result of significant changes in human activity, including population growth, increased demand for animal protein, increased wealth and rapid travel by people and their animals, changes to the environment, and human encroachment on farm land and previously undisturbed wildlife habitats. Other pathogens could follow a similar pathway. … It is very important for policy makers to understand the kind of surveillance and action that will be needed to protect the public and the benefits they provide, and it is up to the scientific and public health community to make this case. … The current status of global surveillance systems in both human and animal populations and the strength of the veterinary health systems are insufficient to preempt a pandemic or to handle an emerging one. "What's needed is a new paradigm, a means for tapping expertise from all sectors, and thinking in a broadly preventive way, to reform animal husbandry and alter the ways people and industries interact with domestic animals and wildlife. The interactions among many factors—from rapid mass transportation to increased consumption of animal protein to wilderness encroachment—have intensified the threat posed by zoonotic diseases. The global community involved with disease surveillance and coordination will be needed to confront this challenge. ...The challenge for controlling emerging diseases with such complex etiology is that no single agency has either the mandate or the capacity to address the entire landscape of zoonotic disease. … The immediate challenge is for the world to support disease monitoring in those areas that need it, while the longer term challenge is to develop greater political will to support a truly global approach to surveillance. ...The most promising approach to sustainable global disease surveillance and response is international collaboration, but large, international organizations are only part of the answer. Sustainable, collaborative work must be based close to where the problems are" (NASEM 2008).*

NASEM (2008) noted that surveillance, defined by the WHO as "the systematic ongoing collection, collation, and analysis of data for public health purposes and the timely dissemination of public health information for assessment and public health response as necessary," had been developing in an ad hoc manner. Important organizations conducting surveillance included the Food and Agriculture Organization of the United Nations (FAO), World Organization for Animal Health (OIE), and the World Health Organization (WHO), in addition to surveillance conducted by countries. In 2006, the joint FAO-OIE-WHO Global Early Warning System (GLEWS) was established to assist in early warning, prevention and control of animal disease threats, including zoonoses (WHO 2006). In 2013, the organization evolved to GLEWS+ to, "to advance from reactive to proactive preparedness and prevention, through joint risk assessment for targeted and timely action" (GLEWS 2013).

LMICs and HICs rely on the WHO Global Influenza Surveillance and Response System (GISRS) (WHO, 2020a). HICs also have some form of additional early response system. GISRS laboratories have become COVID-19 testing centres in many countries. Syndromic and sentinel surveillance have monitored community transmission and geographic spread of influenza-like illness (ILI), acute respiratory infection (ARI), and severe acute respiratory infection (SARI). Continued vigilance is needed to detect the emergence of novel zoonotic viruses affecting humans (WHO 2020b).

Presently, there is a need to standardize and harmonize COVID-19 data collection and reporting across jurisdictions within existing surveillance systems (WHO 2020b).

EARLY WARNING AND RESPONSE DATA SYSTEM: EPI-TRACS

We propose a conceptual early warning and response *Epidemiology Translational Research Action Coordinated System* (Epi-TRACS) inspired by Activity Based Intelligence (ABI) (Figure 1). Quinn (2012) describes ABI as:

> *Activity Based Intelligence is an analysis methodology which rapidly integrates data from multiple traditional data sources and other sources around the interactions of people, events and activities, in order to discover relevant patterns, determine and identify change, and characterize those patterns to drive collection and create decision advantage.*

Attwood (2015) emphasises that "*the enterprise must embrace the opportunities inherent to Big Data while also driving toward a unified intelligence strategy.*" He notes that, "*the primary strategy thus far has been acquisition based, looking to industry and research and development organizations to provide the next best tool and software, rather than addressing the more existential requirement of advancing analytical tradecraft and transforming antiquated intelligence analysis and processing methods.*"

Atwood's (2015) further emphasises that, "*to truly revolutionize and fundamentally change from an individual exploitation process to analysis-based tradecraft, the enterprise needs to harness the potential of Big Data, replacing the methodology of individually exploited pieces of data with an ABI approach. This will enable analysts to focus on hard problems with critical timelines as well as normal day-to-day production activities.*" This is also true in the epidemiology domain, most acutely evident in the case of emerging pathogens and pandemics. In epidemiology, also, *the sharp incline in the amount of data, recent information technology (IT) advances, and the ABI methodology impel significant changes within the traditional intelligence production model of PCPAD (planning and direction, collection, processing and exploitation, analysis and production, and dissemination) (Atwood (2015).*

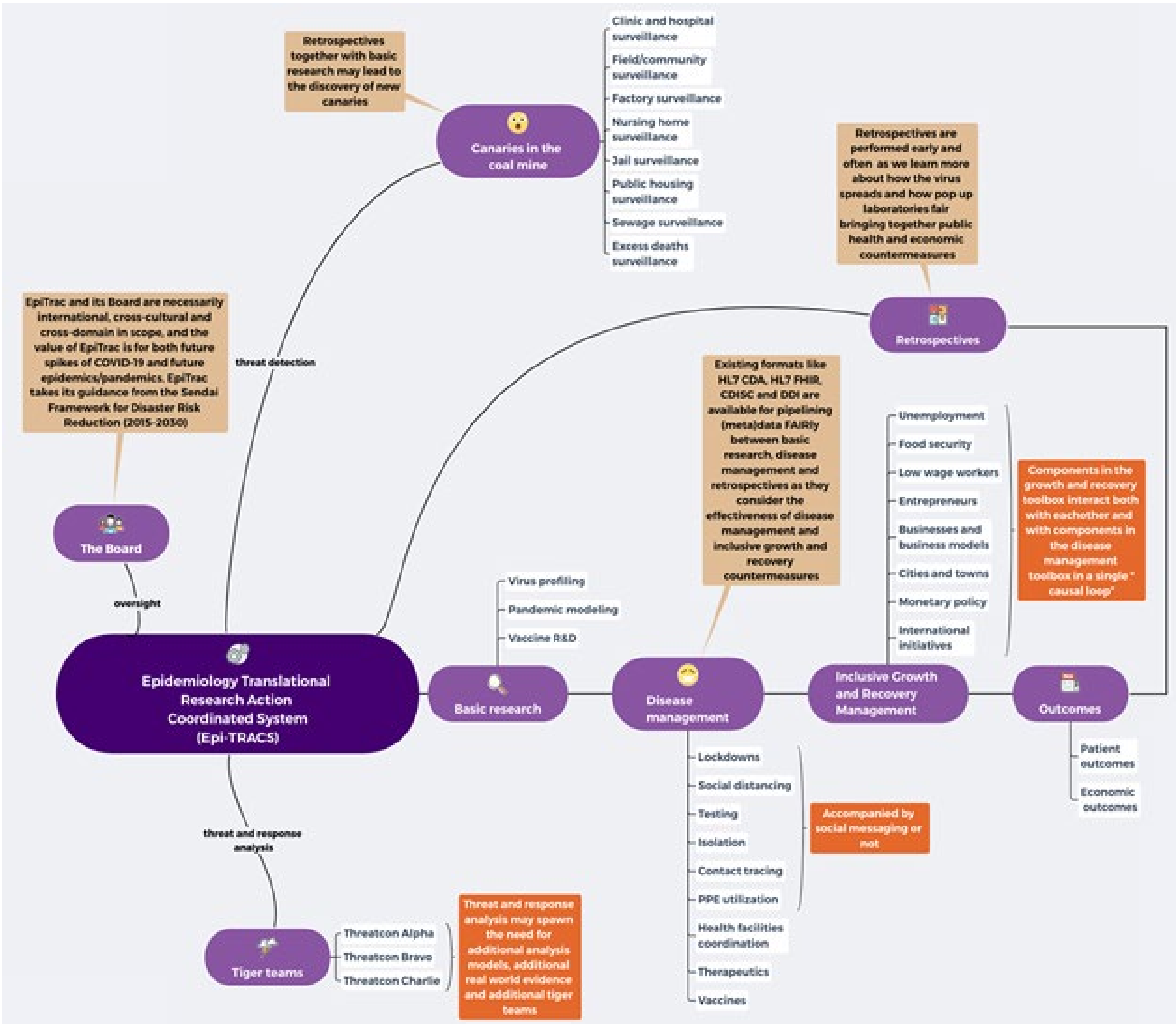

Figure 1. Epi-TRACS. Proposed COVID-19 EPIdemiological Translational Research Action Coordinated System.

Epi-TRACS would use coordinated, rapid response teams for early identification and evaluation of pathogens. To be effective, Epi-TRACS would need to be integrated into existing international surveillance systems across countries and domains. The Epi-TRACS data framework includes sentinel surveillance, public health disease management tools, inclusive growth and recovery tools. Full spectrum epidemiology and the Common Data Model components shown in Figure 1 are discussed elsewhere (Greenfield et al. 2020a,b). The principles of findable, accessible, interoperable, reusable, ethical, and reproducible (FAIRER) data principles and transparency built into Epi-TRACS would contribute to the improving development of decisions and action plans targeting public health and economic outcomes.

AUTHOR ROLES

All authors accept responsibility for the content of the article
Conceptualization: GM, JG, CCA. Methodology: JG, HT, GM. Investigation: CCA. Formal analysis: GM, HT. Writing: JG, CCA. Bibliographic review and analysis: JG, CCA, MC. Visualization: JG, HT, CCA.

Funding: None.
Conflict of interest: None to declare.

REFERENCES

Atwood CP. Activity-Based Intelligence: Revolutionizing Military Intelligence Analysis. Joint Force Quarterly. 2015; 77 (2nd Quarter, April 2015) https://pdfs.semanticscholar.org/b96e/d4e1bcd62fc79eed470db58398d8fbbe514b.pdf?_ga=2.164032895.1918407571.1598743582-256986716.1598743582

Bai Z, Gong Y, Tian X, Cao Y, Liu W, Li J. The Rapid Assessment and Early Warning Models for COVID-19 [published online ahead of print, 2020 Apr 1]. Virol Sin. 2020;1-8. doi:10.1007/s12250-020-00219-0

Gates B. Responding to Covid-19 — A Once-in-a-Century Pandemic? *N Engl J Med* 2020; 382:1677-1679. doi:10.1056/NEJMp2003762

Greenfield J, Nagrani R, Sears M, Mazzaferro G, Austin CC; and the RDA-COVID19-WG. (2020). A Full Spectrum View of the COVID-19 data domain: An Epidemiological Data Framework. In COVID-19 Data sharing in epidemiology, version 0.06b. Research Data Alliance RDA-COVID19-Epidemiology WG. https://doi.org/10.15497/rda00049

Greenfield, J., Sears, M., Nagrani, R., Mazzaferro, G., Widyastuti, A., Austin, C. C.; and the RDA-COVID19-WG. (2020). Common Data Models and Full Spectrum Epidemiology: Epi-STACK framework for COVID-19 epidemiology datasets. In *COVID-19 Data sharing in epidemiology, version 0.06b*. Research Data Alliance RDA-COVID19-Epidemiology WG. https://doi.org/10.15497/rda00049

NASEM (2008). Achieving Sustainable Global Capacity for Surveillance and Response to Emerging Diseases of Zoonotic Origin: Workshop Summary. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/12522

NASEM (2009a). Sustaining Global Surveillance and Response to Emerging Zoonotic Diseases. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/12625

NASEM (2009b). Achieving an Effective Zoonotic Disease Surveillance System. In Sustaining Global Surveillance and Response to Emerging Zoonotic Diseases. National Academies of Sciences, Engineering, and Medicine. https://www.ncbi.nlm.nih.gov/books/NBK215315/

NASEM (2010). Infectious Disease Movement in a Borderless World: Workshop Summary. National Academies of Sciences, Engineering, and Medicine. https://www.nap.edu/download/12758

NASEM (2015). Emerging Viral Diseases: The One Health Connection: Workshop Summary. National Academies of Sciences, Engineering, and Medicine. https://www.nap.edu/catalog/18975/emerging-viral-diseases-the-one-health-connection-workshop-summary

NASEM (2018). Exploring Lessons Learned from Partnerships to Improve Global Health and Safety: Workshop in Brief. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/21690

NASEM (2020a). Rapid Expert Consultation on Data Elements and Systems Design for Modeling and Decision Making for the COVID-19 Pandemic (March 21, 2020). National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25755

NASEM (2020b). Rapid Expert Consultations on the COVID-19 Pandemic: March 14, 2020-April 8, 2020 (National Academies of Sciences, Engineering, and Medicine). National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25784

NASEM (2020c). Evaluating Data Types: A Guide for Decision Makers using Data to Understand the Extent and Spread of COVID-19. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25826

Quinn, Kristin (2012). A Better Toolbox: Analytic Methodology Has Evolved Significantly Since the Cold War, Trajectory (Winter 2012), 1–8. https://trajectorymagazine.com/a-better-toolbox-2/

WHO. [Webpage] Global Influenza Surveillance and Response System (GISRS). https://www.who.int/influenza/gisrs_laboratory/en/. 2020a

WHO. [Webpage] Preparing GISRS for the upcoming influenza seasons during the COVID-19 pandemic – practical considerations. https://apps.who.int/iris/bitstream/handle/10665/332198/WHO-2019-nCoV-Preparing_GISRS-2020.1-eng.pdf. 2020b

DISCUSSION PAPER

ANNEX 6 – **COVID-19 Emergency public health and economic measures causal loop: Laying the groundwork for a computable framework**

Henri E.Z. Tonnang[1], Jay Greenfield[2], Gary Mazzaferro[3], Claire C. Austin[4, §]; and the RDA-COVID19-WG[5]

[1]International Center of Insect Physiology and Ecology (ICIPE), Kenya, [2]Data Documentation Initiative, USA, [3]Consultant, USA, [4]Environment and Climate Change Canada, [5]This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

All views and opinions expressed are those of the co-authors, and do not necessarily reflect the official policy or position of their respective employers, or of any government, agency or organization.

[§] Corresponding authors: htonnang@icipe.org and claire.austin@canada.ca

CITE AS: Henri E.Z. Tonnang[§], Jay Greenfield, Gary Mazzaferro, Claire C. Austin[§]; and the RDA COVID-19 WG (2020). *COVID-19 Emergency public health and economic measures causal loops: A computable framework*. 8 pages. Version 1.0. Available at: http://dx.doi.org/10.2139/ssrn.3686027

ABSTRACT

COVID-19 infections follow dynamic patterns across society. We propose using a "systems thinking" approach to better understand multiple interrelated factors and impacts. Building upon the work of others, we have extended the use of Causal Loop Diagrams (CLD) for whole system qualitative analysis of the linkages and interrelationships between COVID-19 contagion, healthcare, and economic components. The approach used is generic and can easily be applied to both developing and developed countries.

KEYWORDS: Contagion, healthcare, economy, holistic, system thinking, reinforcing and balancing, causal loop diagram (CLD), systems research.

PLEASE SEE THE SSRN PREPRINT FOR THE FULL TEXT (8 pages)

DISCUSSION PAPER

ANNEX 7 – **Common Data Models and Full Spectrum Epidemiology: An Epi-STACK framework for COVID-19 epidemiology datasets**

Jay Greenfield[1, §], Meg Sears[2], Rajini Nagrani[3], Gary Mazzaferro[4], Anna Widyastuti[5], Claire C. Austin[6, §]; and the RDA-COVID19-WG[7]

[1]Data Documentation Initiative, [2]Ottawa Hospital Research Institute, [3]Leibniz Institute for Prevention Research and Epidemiology-BIPS, [4]Consultant, [5] This work was developed as part of the Research Data Alliance RDA-COVID19-WG Recommendations and guidelines on data sharing, and the RDA COVID-19 Epidemiology WG Data sharing in epidemiology, and we acknowledge the support provided by the RDA community and structures.

All views and opinions expressed are those of the co-authors, and do not necessarily reflect the official policy or position of their respective employers, or of any government, agency, or organization.

[§] Corresponding authors: nightcleaner@gmail.com and claire.austin@canada.ca

CITATION: Jay Greenfield, Meg Sears, Rajini Nagrani, Gary Mazzaferro, Anna Widyastuti, Claire C. Austin; and the RDA-COVID19-WG. (2020). *Common Data Models and Full Spectrum Epidemiology: Epi-STACK framework for COVID-19 epidemiology datasets.* [IN: Sharing COVID-19 epidemiology data, Research Data Alliance RDA COVID-19 Epidemiology WG]. https://doi.org/10.15497/rda00049

ABSTRACT

We propose an Epi-STACK framework for hosting COVID-19 datasets compatible with the Common Data Model (CDM). The framework supports full spectrum COVID-19 research, epidemiological and demographic surveillance, clinical care data, and emulated trials.

KEYWORDS: Common Data Model, CDM, epidemiology

BACKGROUND

The World Health Organization (WHO) established an Information Network for Epidemics (EPI-WIN) following the outbreak of COVID-19 disease (WHO 2020). The present paper is the last in a series of four discussion papers leading to the proposal of an Epi-STACK framework based on Common Data Models (CDM). We previously proposed a full-spectrum epidemiology data framework (Greenfield et al. 2020a), Epi-TRACS early warning and response system (Greenfield et al. 2020b), and a computable framework based on systems analysis (Tonnang et al. 2020). The present paper provides an activity-based intelligence framework that could help support development of WHO recommendations and guidelines that are communicated and amplified by the WHO EPI-WIN network.

Clinical and observational research are often conducted as stand-alone studies with little integration between them even when conducted in the same population. The purpose for which clinical data are collected may not meet the data requirements of epidemiology studies. In addition, medical personnel may compromise the quality of data collected when the data are not perceived to be relevant to treatment.

THE COMMON DATA MODEL

Common Data Models (CDMs) have been developed to host and support twenty-first century patient care research using Electronic Health Record (EHR) data. CDMs require standardized/consistent data to facilitate the exchange, pooling, sharing, or storing of data from multiple sources. Within the last decade, several CDMs have been developed collaboratively, and have risen to the level of *de facto* standards for clinical research data (Garza et al. 2016). CDMs may also enable the exploration of observational, quasi-experimental, and experimental clinical studies using healthcare data (Garza et al. 2016). Quasi-experimental designs include the use of EHR data to support emulated clinical trials when clinical trials are impractical (Hernán and Robins, 2016). CDMs support multiple data sources and have evolved to include questionnaire data from field surveys (Blacketer et al. 2015).

Combined data from EHR and survey questionnaires is integral to an epidemiology intelligence framework incorporating person-level clinical and community data to evaluate public health measures. Such an intelligence framework supports cross-domain full spectrum epidemiology (Atwood, 2015).

In parallel with development of CDM approaches, efforts have been made to streamline Case Report Forms (CRFs) and population-based questionnaires (Garza et al. 2016). However, little has been done to integrate the two. We therefore propose the use of the CDM to integrate CRFs and COVID-19 epidemiology data in an Epi-STACK.

In support of Health Care risk assessment research, inputs to the Common Data Model may include health data and community data that take the form of epidemiological surveillance questionnaires administered repeatedly over time (**Figure 1**).

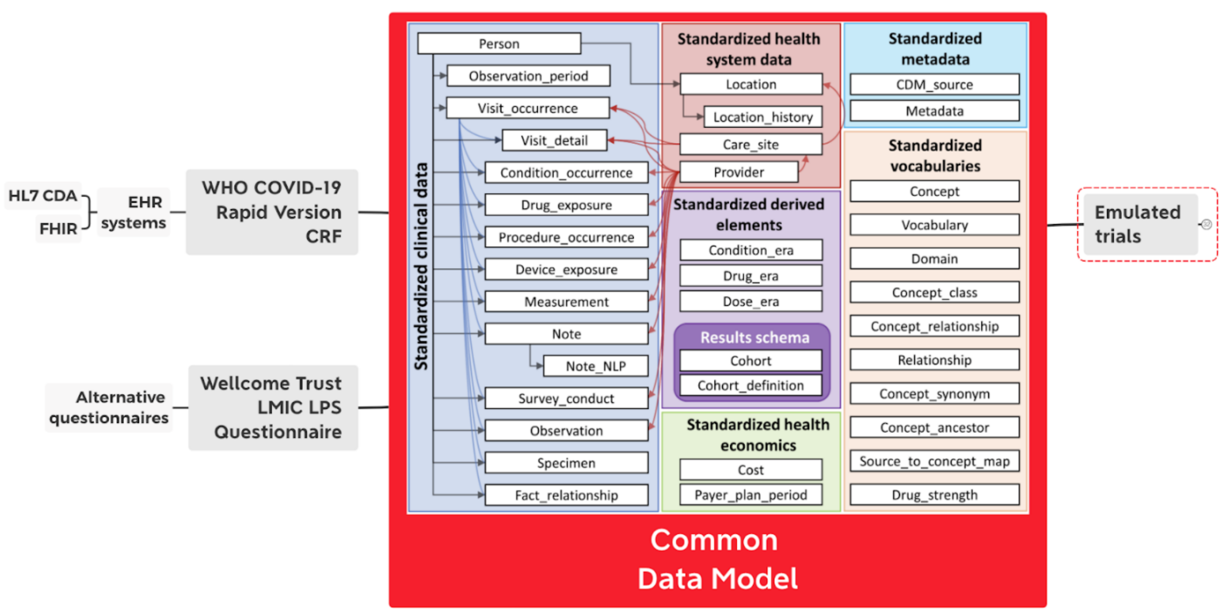

Figure 1. Full spectrum translational research enabled by the Common Data Model. The Common Data Model can be a full spectrum provider for emulated trials. [SOURCE: Book of OHDSI, Chapter 4 (Common Data Model), C. Blacketer, 2020. Creative Commons CC0 v1.0.]

The Common Data Model includes schemas in support of clinical care data and questionnaires, including standardized metadata. Standardized metadata in turn makes full spectrum epidemiology FAIRER (Findable, Accessible, Interoperable, Reusable, Ethical and Reproducible). In fact, several FAIR assessments of a CDM called the Observational Medical Outcomes Partnership (OMOP) have been undertaken (OHDSI 2015). The European Health Data and Evidence Network (EHDEN) initiative tested OMOP for FAIR in a task that required OMOP to harmonize 100 million health records gathered from many sources (van Bochove et al. 2020). A more recent EHDEN initiative has tasked OMOP to go full spectrum in COVID-19 research by combining registry and cohort data alongside EHRs and hospital information (The EHDEN Consortium 2020).

The CDM also supports a risk assessment in emulated trials (**Figures 1 and 2**), where the impact of various disease management and treatments taken separately or together may be assessed using EHR and/or claims data alongside exposure and/or other confounding/stratification variables (Gershman et al. 2018).

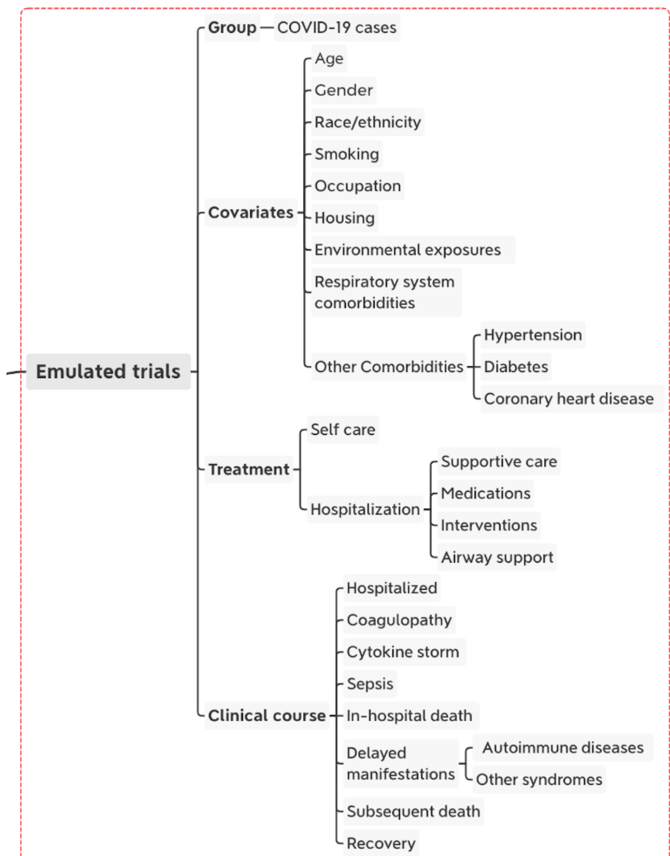

Figure 2. COVID-19 Risk Assessment for Health Care and Population Health.

EPI-STACK

An activity-based intelligence framework was adapted to create Epi-STACK to support cross-domain, full spectrum COVID-19 epidemiology (**Figure 3**). Epi-STACK may support many types of tools and analytical models surrounding pandemic response. Emulated trials can be explored to discern associations and causal relationships, as illustrated in **Figure 2**. The COVID-19 *Risk Assessment Research Framework* is shown in **Figure 3**.

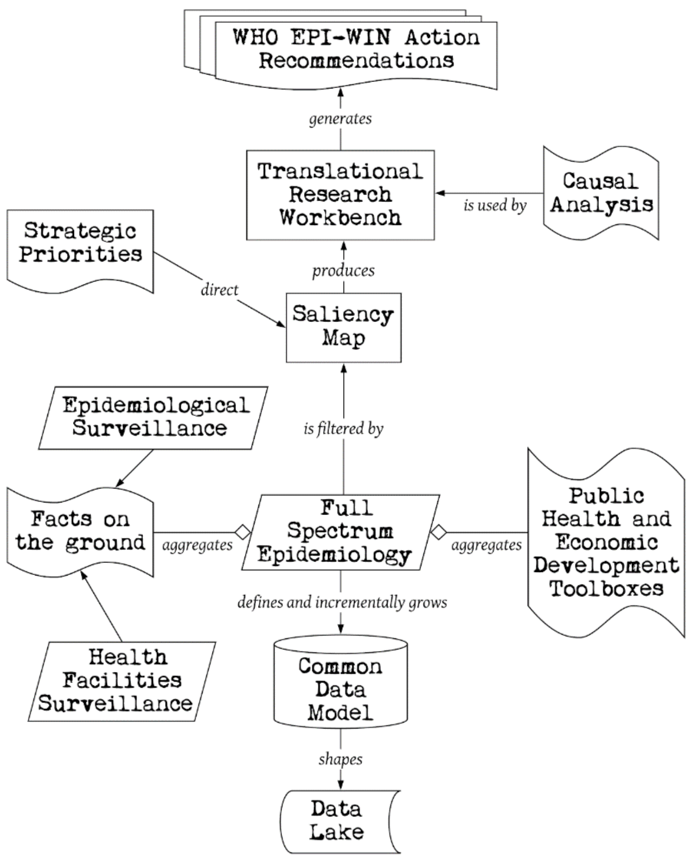

Figure 3. Epi-STACK: A framework inspired by Activity Based Intelligence to support full spectrum COVID-19 epidemiology.

Full spectrum epidemiology considers multiple data types disease surveillance sources together with the public health and economic development that mitigate spread and severity of disease in a population **(Figure 2)**. The proposed full spectrum COVID-19 epidemiology **(Figure 2)** would be hosted by a CDM. One of the many research scenarios that Epi-STACK could support would be a *Risk Assessment Research Framework* for COVID-19 Health Care using causal analysis in quasi-experimental designs **(Figure 3)**.

Epi-STACK and EPI-WIN

The CDM is the starting point of a customized workflow to host full spectrum epidemiology (**Figure 4**).

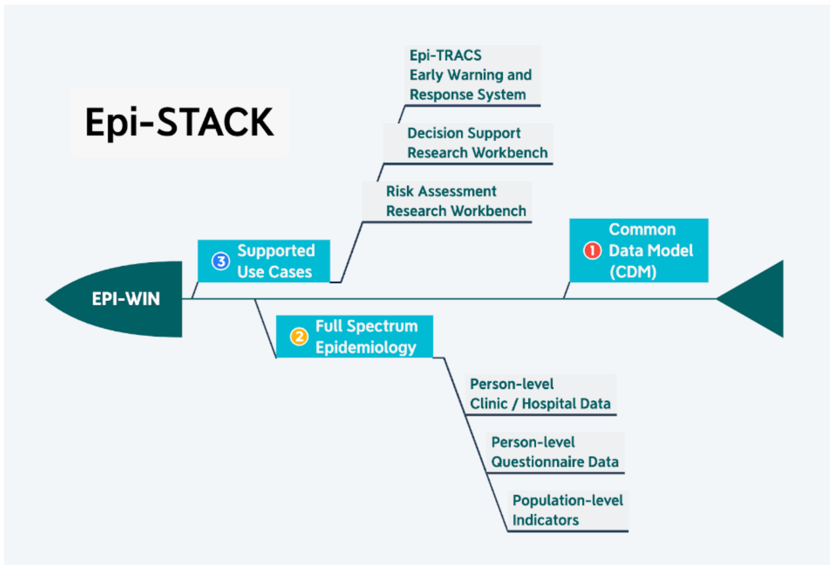

Figure 4. Epi-STACK showing three use cases.

Risk assessment using emulated trials is one use case for the Common Data Model. Other use cases could include an early warning and response system, decision support systems for public health and economic interventions (Tonnang et al. 2020), as well as Epi-TRACS (Greenfield et al. 2020). These use cases define our proposed Full Spectrum Epidemiology, which in turn gives shape to the CDM. All of these use cases require a time series of clinical datasets, survey datasets, and population indicators which make Epi-STACK fit for purpose. Epi-STACK has the ability over time to produce a series of interim results that WHO Epi-WIN would be able to channel to a wide range of target audiences (**Figure 4**).

Author roles

All authors accept responsibility for the content of the article

Conceptualization: GM MS RN CCA. Methodology: GM HT. Investigation: MC. Formal analysis: CCA. Validation: CCA MC. Writing: JG MS CCA. Bibliographic review and analysis: MC AW. Visualization: JG

Funding: None.

Conflict of interest: None to declare.

REFERENCES

Atwood CP. Activity-Based Intelligence: Revolutionizing Military Intelligence Analysis. *Joint Force Quarterly.* 2015; *77 (*2nd Quarter, April 2015)

Blacketer M, Voss EA, Ryan PB. Applying the OMOP Common Data Model to Survey Data [Webpage]. 2015 https://www.ohdsi.org/web/wiki/lib/exe/fetch.php?media=resources:using_the_omop_cdm_with_survey_and_registry_data_v6.0.pdf.

Docherty AB, Harrison EM, Green CA, Hardwick HE, Pius R, Norman L, et al. Features of 20 133 UK patients in hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: prospective observational cohort study. BMJ [Internet]. 2020 May 22 [cited 2020 Jul 8];369. Available from: https://www.bmj.com/content/369/bmj.m1985Garza M, Fiol GD, Tenenbaum J, et al. Evaluating common data models for use with a longitudinal community registry. *Journal of Biomedical Informatics*. 2016;64(12):333-341. DOI: https://doi.org/10.1016/j.jbi.2016.10.016

Gershman B, Guo DP, Dahabreh IJ. Using observational data for personalized medicine when clinical trial evidence is limited. *Fertil Steril*. 2018;109(6):946-951. doi:10.1016/j.fertnstert.2018.04.005

Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *Am J Epidemiol*. 2016;183(8):758-764. doi:10.1093/aje/kwv254

OHDSI (2015). OMOP Common Data Model. Observational Health Data Sciences and Informatics Collaborative http://www.ohdsi.org/data-stohdsiandardization/the-common-data-model/, 20 Jul 2015.

OHDSI (2020). The Book of OHDSI. Observational Health Data Sciences and Informatics Collaborative. https://ohdsi.github.io/TheBookOfOhdsi/

The EHDEN Consortium. Data Partner Rapid Collaboration Call on COVID-19. V 1.4, April 15 2020

van Bochove K, Vos E, van Winzu A, Kurps J, Moinat M. Implementing FAIR in OHDSI and EHDEN. https://www.ohdsi.org/wp-content/uploads/2020/05/Implementing-FAIR-in-OHDSI.pdf. The Hyve, Utrecht, The Netherlands. 2020

WHO. EPI-WIN: WHO information network for epidemics [Webpage]. 2020; https://www.who.int/teams/risk-communication

RESEARCH DATA ALLIANCE

RDA COVID-19 Epidemiology WG

Bibliography

Abd-Alrazaq, A., Alhuwail, D., Househ, M., Hamdi, M., & Shah, Z. (2020). Top Concerns of Tweeters During the COVID-19 Pandemic: Infoveillance Study. *JOURNAL OF MEDICAL INTERNET RESEARCH*, *22*(4). https://doi.org/10.2196/19016

Abdulmajeed, K., Adeleke, M., & Popoola, L. (2020). Online forecasting of COVID-19 cases in Nigeria using limited data. *Data in Brief*, 105683. https://doi.org/10.1016/j.dib.2020.105683

Access Now. (2020). *Recommendations on privacy and data protection in the fight against COVID-19*. https://www.accessnow.org/cms/assets/uploads/2020/03/Access-Now-recommendations-on-Covid-and-data-protection-and-privacy.pdf

Aguilar-Gallegos, N., Romero-García, L. E., Martínez-González, E. G., García-Sánchez, E. I., & Aguilar-Ávila, J. (2020). Dataset on dynamics of Coronavirus on Twitter. *Data in Brief*, 105684. https://doi.org/10.1016/j.dib.2020.105684

Ahn, M., Anderson, D. E., Zhang, Q., Tan, C. W., Lim, B. L., Luko, K., Wen, M., Chia, W. N., Mani, S., Wang, L. C., Ng, J. H. J., Sobota, R. M., Dutertre, C.-A., Ginhoux, F., Shi, Z.-L., Irving, A. T., & Wang, L.-F. (2019). Dampened NLRP3-mediated inflammation in bats and implications for a special viral reservoir host. *Nature Microbiology*, *4*(5), 789–799. https://doi.org/10.1038/s41564-019-0371-3

AI challenge with AI2, CZI, MSR, Georgetown, NIH, & The White House. (2020). *COVID-19 Open Research Dataset Challenge (CORD-19)*. https://kaggle.com/allen-institute-for-ai/CORD-19-research-challenge

Aitsi-Selmi, A., & Murray, V. (2016). Protecting the Health and Well-being of Populations from Disasters: Health and Health Care in The Sendai Framework for Disaster Risk Reduction 2015-2030. *Prehospital and Disaster Medicine*, *31*(1), 74–78. Cambridge Core. https://doi.org/10.1017/S1049023X15005531

Allen, T., Murray, K. A., Zambrana-Torrelio, C., Morse, S. S., Rondinini, C., Di Marco, M., Breit, N., Olival, K. J., & Daszak, P. (2017). Global hotspots and correlates of emerging zoonotic diseases. *Nature Communications*, *8*(1). Scopus. https://doi.org/10.1038/s41467-017-00923-8

Allocati, N., Petrucci, A. G., Di Giovanni, P., Masulli, M., Di Ilio, C., & De Laurenzi, V. (2016). Bat–man disease transmission: Zoonotic pathogens from wildlife reservoirs to human populations. *Cell Death Discovery*, *2*(1), 1–8. https://doi.org/10.1038/cddiscovery.2016.48

Andersen, K. G., Rambaut, A., Lipkin, W. I., Holmes, E. C., & Garry, R. F. (2020). The proximal origin of SARS-CoV-2. *Nature Medicine*, *26*(4), 450–452. https://doi.org/10.1038/s41591-020-0820-9

Anderson, R. M., Fraser, C., Ghani, A. C., Donnelly, C. A., Riley, S., Ferguson, N. M., Leung, G. M., Lam, T. H., & Hedley, A. J. (2004). Epidemiology, transmission dynamics and control of SARS: The 2002-2003 epidemic. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *359*(1447), 1091–1105. https://doi.org/10.1098/rstb.2004.1490

Angelopoulos, A. N., Pathak, R., Varma, R., & Jordan, M. I. (2020). On Identifying and Mitigating Bias in the Estimation of the COVID-19 Case Fatality Rate. *Harvard Data Science Review*. https://hdsr.mitpress.mit.edu/pub/y9vc2u36/release/2

Apple. (2020). *COVID-19—Mobility Trends Reports*. AppleMaps. https://www.apple.com/covid19/mobility

Bahri, M. (2020). *The Nexus Impacts of the COVID-19: A Qualitative Perspective*. https://doi.org/10.20944/preprints202005.0033.v1

Bai, Z., Gong, Y., Tian, X., Cao, Y., Liu, W., & Li, J. (2020). The Rapid Assessment and Early Warning Models for COVID-19. *Virologica Sinica*, 1–8. https://doi.org/10.1007/s12250-020-00219-0

Bajwa, S. J. S., Sarna, R., Bawa, C., & Mehdiratta, L. (2020). Peri-operative and critical care concerns in coronavirus pandemic. *INDIAN JOURNAL OF ANAESTHESIA*, *64*(4), 267–274. https://doi.org/10.4103/ija.IJA_272_20

Barton, C. M., Alberti, M., Ames, D., Atkinson, J.-A., Bales, J., Burke, E., Chen, M., Diallo, S. Y., Earn, D. J. D., Fath, B., Feng, Z., Gibbons, C., Hammond, R., Heffernan, J., Houser, H., Hovmand, P. S., Kopainsky, B., Mabry, P. L., Mair, C., … Tucker, G. (2020). Call for transparency of COVID-19 models. *Science*, *368*(6490), 482.2-483. https://doi.org/10.1126/science.abb8637

Battegay, M., Kuehl, R., Tschudin-Sutter, S., Hirsch, H. H., Widmer, A. F., & Neher, R. A. (2020). 2019-novel Coronavirus (2019-nCoV): Estimating the case fatality rate – a word of caution. *Swiss Medical Weekly*, *150*(0506). https://doi.org/10.4414/smw.2020.20203

Berry, I., Soucy, J.-P. R., Tuite, A., & Fisman, D. (2020). Open access epidemiologic data and an interactive dashboard to monitor the COVID-19 outbreak in Canada. *Canadian Medical Association Journal*, *192*(15), E420–E420. https://doi.org/10.1503/cmaj.75262

Blacketer, M. S., Voss, E. A., & Ryan, P. B. (2013). Applying the OMOP Common Data Model to Survey Data. *Medical Care*, S45–S52. https://www.ohdsi.org/web/wiki/lib/exe/fetch.php?media=resources:using_the_omop_cdm_with_survey_and_registry_data_v6.0.pdf.

Bradley, D. T., Mansouri, M. A., Kee, F., & Garcia, L. M. T. (2020). A systems approach to preventing and responding to COVID-19. *EClinicalMedicine*, *21*. https://doi.org/10.1016/j.eclinm.2020.100325

Brickley, D., et al., & et al. (2020). *Organization of Schemas*. https://schema.org/docs/schemas.html

Bullock, H. E., Harlow, L. L., & Mulaik, S. A. (2009). Causation Issues in Structural Equation Modeling Research. *Structural Equation Modeling: A Multidisciplinary Journal*, *1*(3), 253–267. Scopus. https://doi.org/10.1080/10705519409539977

Burke, R. L., Kronmann, K. C., Daniels, C. C., Meyers, M., Byarugaba, D. K., Dueger, E., Klein, T. A., Evans, B. P., & Vest, K. G. (2012). A Review of Zoonotic Disease Surveillance Supported by the Armed Forces Health Surveillance Center. *Zoonoses and Public Health*, *59*(3), 164–175. Scopus. https://doi.org/10.1111/j.1863-2378.2011.01440.x

Calisher, C. H., Childs, J. E., Field, H. E., Holmes, K. V., & Schountz, T. (2006). Bats: Important Reservoir Hosts of Emerging Viruses. *Clinical Microbiology Reviews*, *19*(3), 531–545. https://doi.org/10.1128/CMR.00017-06

CDC. (2018, December 4). *National Pandemic Strategy*. https://www.cdc.gov/flu/pandemic-resources/national-strategy/index.html

CDC. (2020a). *Cases of COVID19 in the U.S.* [Dataset]. Centers for Disease Control, USA. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/cases-in-us.html

CDC. (2020b). *Weekly provisional death counts by select demographic and geographic characteristics* [Dataset]. Center for Disease Control, USA. https://www.cdc.gov/nchs/nvss/vsrr/covid_weekly/index.htm

CDC. (2020c). *Weekly provisional death counts from death certificate data: COVID19, pneumonia, flu* [Dataset]. Center for Disease Control, USA. https://www.cdc.gov/nchs/nvss/vsrr/covid19/index.htm

CDC. (2020d, February 19). *Zoonotic Diseases—One Health*. https://www.cdc.gov/onehealth/basics/zoonotic-diseases.html

CDC. (2020e, March 24). *Public Health and Promoting Interoperability Programs (formerly, known as Electronic Health Records Meaningful Use)*. https://www.cdc.gov/ehrmeaningfuluse/introduction.html

CDC. (2020f). *Person Under Investigation (PUI) and Case Report F.pdf*. 2. https://www.phenxtoolkit.org/toolkit_content/PDF/CDC_PUI.pdf

CDISC. (2020). *CDISC Standards in the Clinical Research Process* [Text]. CDISC. https://www.cdisc.org/standards

CGH. (2019, January 16). *Global Health Security Agenda: GHSA Zoonotic Disease Action Package (GHSA Action Package Prevent-2)*. Center for Global Health. https://www.cdc.gov/globalhealth/security/actionpackages/zoonotic_disease.htm

Chan, A. T., Drew, D. A., Nguyen, L. H., Joshi, A. D., Ma, W., Guo, C.-G., Lo, C.-H., Mehta, R. S., Kwon, S., Sikavi, D. R., Magicheva-Gupta, M. V., Fatehi, Z. S., Flynn, J. J., Leonardo, B. M., Albert, C. M., Andreotti, G., Beane Freeman, L. E., Balasubramanian, B. A., Brownstein, J. S., … Spector, T. (2020). The Coronavirus Pandemic Epidemiology (COPE) Consortium: A Call to Action. *Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology*. https://doi.org/10.1158/1055-9965.EPI-20-0606

Chan, J. F.-W., To, K. K.-W., Tse, H., Jin, D.-Y., & Yuen, K.-Y. (2013). Interspecies transmission and emergence of novel viruses: Lessons from bats and birds. *Trends in Microbiology*, *21*(10), 544–555. https://doi.org/10.1016/j.tim.2013.05.005

Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., & Zhang, L. (2020). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *The Lancet*, *395*(10223), 507–513. https://doi.org/10.1016/S0140-6736(20)30211-7

Cheng, V. C. C., Wong, S.-C., Chen, J. H. K., Yip, C. C. Y., Chuang, V. W. M., Tsang, O. T. Y., Sridhar, S., Chan, J. F. W., Ho, P.-L., & Yuen, K.-Y. (2020). Escalating infection control response to the rapidly evolving epidemiology of the coronavirus disease 2019 (COVID-19) due to SARS-CoV-2 in Hong Kong. *INFECTION CONTROL AND HOSPITAL EPIDEMIOLOGY*, *41*(5), 493–498. https://doi.org/10.1017/ice.2020.58

Chowdhury, R., Heng, K., Shawon, M. S. R., Goh, G., Okonofua, D., Ochoa-Rosales, C., Gonzalez-Jaramillo, V., Bhuiya, A., Reidpath, D., Prathapan, S., Shahzad, S., Althaus, C. L., Gonzalez-Jaramillo, N., Franco, O. H., & The Global Dynamic Interventions Strategies for COVID-19 Collaborative Group. (2020). Dynamic interventions to control COVID-19 pandemic: A multivariate prediction modelling study comparing 16 worldwide countries. *European Journal of Epidemiology*. https://doi.org/10.1007/s10654-020-00649-w

COAR. (2020). *Recommendations for COVID-19 resources in repositories*. Confederation of Open Access Repositories. https://www.coar-repositories.org/news-updates/covid19-recommendations/

Coccia, M. (2020). *Two mechanisms for accelerated diffusion of COVID-19 outbreaks in regions with high intensity of population and polluting industrialization: The air pollution-to-human and human-to-human transmission dynamics*. Cold Spring Harbor Laboratory Press. https://www.medrxiv.org/content/10.1101/2020.04.06.20055657v1

CODATA, C. on D. of the I. S. C., CODATA International Data Policy Committee, CODATA and CODATA China High-level International Meeting on Open Research Data Policy and Practice, Hodson, S., Mons, B., Uhlir, P., & Zhang, L. (2019). *The Beijing Declaration on Research Data*. https://doi.org/10.5281/zenodo.3552330

CODATA, R. (2017). *International Research Data Management glossary (IRiDiuM)*. https://codata.org/initiatives/working-groups/standard-glossary-for-research-data-management-iridium/

COVID19India. (2020). *Coronavirus in India: Latest Map and Case Count*. https://www.covid19india.org

Datacovid. (2020). *Barometer Covid19*. https://datacovid.org/

Davis, L. (2020). *Corona Data Scraper* [HTML]. COVID Atlas. https://github.com/covidatlas/coronadatascraper (Original work published 2020)

Day, M. J., Breitschwerdt, E., Cleaveland, S., Karkare, U., Khanna, C., Kirpensteijn, J., Kuiken, T., Lappin, M. R., McQuiston, J., Mumford, E., Myers, T., Palatnik-de-Sousa, C. B., Rubin, C., Takashima, G., & Thiermann, A. (2012). *Surveillance of Zoonotic Infectious Disease Transmitted by Small Companion Animals—Volume 18, Number 12—December 2012—Emerging Infectious Diseases journal—CDC*. https://doi.org/10.3201/eid1812.120664

DDI Alliance. (2020). *Data Documentation Initiative*. https://ddialliance.org/

De Silva, D. S., Galobardes, B., Plummer, J., Herbst, E., Todd, J., Kanjala, C., Le Doare, L. D., Stewart, R., Mridha, M., Novella, R., Slaymaker, E., Crampin, M., Ramsay, M., Ziraba, A., & the LMIC Working Group. (2020, May). *LMIC Covid Core Questionnaire*. https://wellcomecloud-my.sharepoint.com/:w:/g/personal/b_galobardes_wellcome_ac_uk/EX963zhZHWlPkYsDM79jQcwBoK4RD7JrYx8k4YFo-Ep6mA?rtime=k2_HF54b2Eg

DeBord, D. G., Carreon, T., & Lentz, T. (2016). Use of the "Exposome" in the Practice of Epidemiology: A Primer on -Omic Technologies. *Am J Epidemiology*, *184*(4), 312–314. https://doi.org/10.1093/aje/kwv325

Degeling, C., Johnson, J., Ward, M., Wilson, A., & Gilbert, G. (2017). A Delphi Survey and Analysis of Expert Perspectives on One Health in Australia. *EcoHealth*, *14*(4), 783–792. Scopus. https://doi.org/10.1007/s10393-017-1264-7

Djalante, R., Shaw, R., & DeWit, A. (2020). Building resilience against biological hazards and pandemics: COVID-19 and its implications for the Sendai Framework. *Progress in Disaster Science*, *6*, 100080. https://doi.org/10.1016/j.pdisas.2020.100080

Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet. Infectious Diseases*, *20*(5), 533–534. PubMed. https://doi.org/10.1016/S1473-3099(20)30120-1

Drew, D. A., Nguyen, L. H., Steves, C. J., Menni, C., Freydin, M., Varsavsky, T., Sudre, C. H., Cardoso, M. J., Ourselin, S., Wolf, J., Spector, T. D., Chan, A. T., & COPE Consortium. (2020). Rapid implementation of mobile technology for real-time epidemiology of COVID-19. *Science (New York, N.Y.)*. https://doi.org/10.1126/science.abc0473

DSI, D. S., CEF eHealth. (2020, March 24). *EHDSI INTEROPERABILITY SPECIFICATIONS, Requirements and Frameworks (normative artefacts)—EHealth DSI Operations—CEF Digital*. https://ec.europa.eu/cefdigital/wiki/pages/viewpage.action?pageId=35210463

Duncan, G. T., Elliot, M., & Salazar, G. J. J. (2011). *Statistical Confidentiality: Principles and Practice*. Springer-Verlag. https://doi.org/10.1007/978-1-4419-7802-8

Duncan, M. A., Drociuk, D., Belflower-Thomas, A., Van Sickle, D., Gibson, J. J., Youngblood, C., & Daley, W. R. (2011). Follow-Up Assessment of Health Consequences after a Chlorine Release from a Train Derailment-Graniteville, SC, 2005. *Journal of Medical Toxicology*, *7*(1), 85–91. Scopus. https://doi.org/10.1007/s13181-010-0130-6

eCDC. (2015). *Country preparedness plans on zoonotic influenza*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/avian-influenza-humans/country-preparedness-plans-avian-influenza-humans

ECDC. (2019). *The European Surveillance System (TESSy)*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/publications-data/european-surveillance-system-tessy

eCDC. (2019, December 12). *The European Union One Health 2018 Zoonoses Report*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/publications-data/european-union-one-health-2018-zoonoses-report

ECDC. (2020a). *COVID-19 situation update worldwide, as of 9 June 2020*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

ECDC. (2020b). *How ECDC collects and processes COVID-19 data*. European Centre for Disease Prevention and Control. https://www.ecdc.europa.eu/en/covid-19/data-collection

ECDC. (2020c, April 15). *Geographic distribution of COVID-19 cases worldwide*. https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

ESIP. (2019). *Data Citation Guidelines for Earth Science Data , Version 2*. Earth Science Information Partners. https://doi.org/10.6084/m9.figshare.8441816.v1

EU. (2019). *EHealth Network Guidelines to the EU Member States and the European Commission on an interoperable eco-system for digital health and investment programmes for a new/updated generation of digital infrastructure in Europe, ev_20190611_co922_en.pdf*. EU EHealth Network. https://ec.europa.eu/health/sites/health/files/ehealth/docs/ev_20190611_co922_en.pdf

European Commission. (2020). *Pseudonymisation tool*. EUPID - European Platform on Rare Disease Registration. https://eu-rd-platform.jrc.ec.europa.eu/node/2_en

COMMISSION RECOMMENDATION (EU) 2020/518 of 8 April 2020 on a common Union toolbox for the use of technology and data to combat and exit from the COVID-19 crisis, in particular concerning mobile applications and the use of anonymised mobility data, (2020). https://ec.europa.eu/info/sites/info/files/recommendation_on_apps_for_contact_tracing_4.pdf

Everard, M., Johnston, P., Santillo, D., & Staddon, C. (2020). The role of ecosystems in mitigation and management of Covid-19 and other zoonoses. *Environmental Science and Policy*, *111*, 7–17. https://doi.org/10.1016/j.envsci.2020.05.017

Exaptive. (2020). *Cognitive City: COVID-19 Resources*. Exaptive, and the Bill and Melinda Gates Foundation. https://covid-19.cognitive.city/cognitive/community/resource-gallery

FAIR4Health. (2020). *FAIR4Health at RDA Germany Conference 2020—Resources*. FAIR4Health Consortium. https://www.fair4health.eu/en/resources

Falzon, L. C., Alumasa, L., Amanya, F., Kang'ethe, E., Kariuki, S., Momanyi, K., Muinde, P., Murungi, M. K., Njoroge, S. M., Ogendo, A., Ogola, J., Rushton, J., Woolhouse, M. E. J., & Fèvre, E. M. (2019). One Health in Action: Operational Aspects of an Integrated Surveillance System for Zoonoses in Western Kenya. *Frontiers in Veterinary Science*, *6*. Scopus. https://doi.org/10.3389/fvets.2019.00252

Fauci, A. S., & Morens, D. M. (2012). The Perpetual Challenge of Infectious Diseases. *New England Journal of Medicine*, *366*(5), 454–461. https://doi.org/10.1056/NEJMra1108296
FDA. (2019). *Sentinel Common Data Model | Sentinel Initiative*. FDA Sentinel Initiative. https://www.sentinelinitiative.org/sentinel/data/distributed-database-common-data-model
Ferguson, N. M., Cummings, D. A. T., Fraser, C., Cajka, J. C., Cooley, P. C., & Burke, D. S. (2006). Strategies for mitigating an influenza pandemic. *Nature*, *442*(7101), 448–452. Scopus. https://doi.org/10.1038/nature04795
Fineberg, H. V. (2014). Global health: Pandemic preparedness and response—Lessons from the H1N1 influenza of 2009. *New England Journal of Medicine*, *370*(14), 1335–1342. https://doi.org/10.1056/NEJMra1208802
Finnie, T., South, A., & Bento, A. (2016). EpiJSON: A unified data-format for epidemiology. *Epidemics*, *15*(June, 2016), 20–26. https://doi.org/10.1016/j.epidem.2015.12.002
FitzHenry, F., Resnic, F. S., Robbins, S. L., Denton, J., Nookala, L., Meeker, D., Ohno-Machado, L., & Matheny, M. E. (2015). Creating a Common Data Model for Comparative Effectiveness with the Observational Medical Outcomes Partnership. *Applied Clinical Informatics*, *6*(3), 536–547. https://doi.org/10.4338/ACI-2014-12-CR-0121
Foley, D., Rueda, P., Wilkerson, R., & the ESWG (Entomological Surveillance Working Group). (2019). *VectorMap: Know the vector, know the threat*. Walter Reed Biosystematics Unit (WRBU), U.S. Department of Defense. http://vectormap.si.edu/Project_ESWG.htm
Franklin, A. B., & Bevins, S. N. (2020). Spillover of SARS-CoV-2 into novel wild hosts in North America: A conceptual model for perpetuation of the pathogen. *Science of the Total Environment*, *733*. Scopus. https://doi.org/10.1016/j.scitotenv.2020.139358
French COVID-19. (2020, March 12). COVID-19 funding opportunities. *Reacting*. https://reacting.inserm.fr/covid-19-funding-opportunities/
Frontera, A., Martin, C., Vlachos, K., & Sgubin, G. (2020). Regional air pollution persistence links to COVID-19 infection zoning. *The Journal of Infection*. https://doi.org/10.1016/j.jinf.2020.03.045
Gates, B. (2020). Responding to Covid-19—A once-in-a-century pandemic? *New England Journal of Medicine*, *382*(18), 1677–1679. Scopus. https://doi.org/10.1056/NEJMp2003762
Gebreyes, W. A., Dupouy-Camet, J., Newport, M. J., Oliveira, C. J. B., Schlesinger, L. S., Saif, Y. M., Kariuki, S., Saif, L. J., Saville, W., Wittum, T., Hoet, A., Quessy, S., Kazwala, R., Tekola, B., Shryock, T., Bisesi, M., Patchanee, P., Boonmar, S., & King, L. J. (2014). The Global One Health Paradigm: Challenges and Opportunities for Tackling Infectious Diseases at the Human, Animal, and Environment Interface in Low-Resource Settings. *PLoS Neglected Tropical Diseases*, *8*(11). https://doi.org/10.1371/journal.pntd.0003257
German Data Forum (RatSWD). (2020). Remote Access to data from official statistics agencies and social security agencies. *RatSWD Output Paper Series*. https://www.ratswd.de/en/publication/output-series/2855
German National Cohort. (2020, March 31). *NAKO Gesundheitsstudie—Kontakt*. https://nako.de/allgemeines/kontakt/
Gershman, B., Guo, D. P., & Dahabreh, I. J. (2018). Using observational data for personalized medicine when clinical trial evidence is limited. *Fertility and Sterility*, *109*(6), 946–951. https://doi.org/10.1016/j.fertnstert.2018.04.005

GESIS Panel Team. (2020). *GESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in GermanyGESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in Germany* (1.1.0) [Data set]. GESIS Data Archive. https://doi.org/10.4232/1.13520

Ghani, A. C., Donnelly, C. A., Cox, D. R., Griffin, J. T., Fraser, C., Lam, T. H., Ho, L. M., Chan, W. S., Anderson, R. M., Hedley, A. J., & Leung, G. M. (2005). Methods for Estimating the Case Fatality Ratio for a Novel, Emerging Infectious Disease. *American Journal of Epidemiology*, *162*(5), 479–486. https://doi.org/10.1093/aje/kwi230

GHRU. (2020, March 1). *GHRU - COVID Questionnaire—V6.docx*. Dropbox. https://www.dropbox.com/s/auvvey4utibd85s/GHRU%20-%20COVID%20Questionnaire%20-%20v6.docx?dl=0

Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Filippo, A. D., Matteo, A. D., & Colaneri, M. (2020). Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*, 1–6. https://doi.org/10.1038/s41591-020-0883-7

GLEWS. (2006). *The Global Early Warning System for health threats and emerging risks at the human–animal–ecosystems interface*. http://www.glews.net/

GLOPID, Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2018). *Principles of data sharing in public health emergencies*. https://www.glopid-r.org/wp-content/uploads/2018/06/glopid-r-principles-of-data-sharing-in-public-health-emergencies.pdf

Goldberg, M., & Villeneuve, P. (2020). *Air pollution, COVID-19 and death: The perils of bypassing peer review*. The Conversation. http://theconversation.com/air-pollution-covid-19-and-death-the-perils-of-bypassing-peer-review-136376

Google. (2020a). *Google Dataset Search*. https://datasetsearch.research.google.com/

Google. (2020b). *Google LLC 'COVID-19 Community Mobility Report'*. COVID-19 Community Mobility Report. https://www.google.com/covid19/mobility?hl=en

Gostin, L. O., & Katz, R. (2016). The International Health Regulations: The Governing Framework for Global Health Security. *Milbank Quarterly*, *94*(2), 264–313. https://doi.org/10.1111/1468-0009.12186

Government of the Democratic Republic of the Congo, & World Health Organization. (2019). *Strategic_response_plan.pdf*. https://www.un.org/ebolaresponsedrc/sites/www.un.org.ebolaresponsedrc/files/strategic_response_plan.pdf

Grange, E. S., Neil, E. J., Stoffel, M., Singh, A. P., Tseng, E., Resco-Summers, K., Fellner, B. J., Lynch, J. B., Mathias, P. C., Mauritz-Miller, K., Sutton, P. R., & Leu, M. G. (2020). Responding to COVID-19: The UW Medicine Information Technology Services Experience. *APPLIED CLINICAL INFORMATICS*, *11*(2), 265–275. https://doi.org/10.1055/s-0040-1709715

Haesler, B., Gilbert, W., Jones, B. A., Pfeiffer, D. U., Rushton, J., & Otte, M. J. (2013). The Economic Value of One Health in Relation to the Mitigation of Zoonotic Disease Risks. In Mackenzie, JS and Jeggo, M and Daszak, P and Richt, JA (Ed.), *ONE HEALTH: THE HUMAN-ANIMAL-ENVIRONMENT INTERFACES IN EMERGING INFECTIOUS DISEASES: THE CONCEPT AND EXAMPLES OF A ONE HEALTH APPROACH* (Vol. 365, pp. 127–151). SPRINGER-VERLAG BERLIN. https://doi.org/10.1007/82_2012_239

Hare, S. S., Rodrigues, J. C. L., Jacob, J., Edey, A., Devaraj, A., Johnstone, A., McStay, R., Nair, A., & Robinson, G. (2020). A UK-wide British Society of Thoracic Imaging COVID-19 imaging repository and database: Design, rationale and implications for education and research. *Clinical Radiology*, *75*(5), 326–328. https://doi.org/10.1016/j.crad.2020.03.005

Harvard. (2020). *COVID-19 Hospital Capacity Estimates 2020*. Harvard Global Health Institute. https://globalepidemics.org/

HCSRN. (2019). *VDW Data Model*. Healthcare Systems Research Network. http://www.hcsrn.org/en/Tools%20&%20Materials/VDW/

Healy, K. (2020). *Rpackage (covdata)—COVID19 Case and Mortality Time Series*. https://kjhealy.github.io/covdata

Hernán, M. A., & Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology*, *183*(8), 758–764. https://doi.org/10.1093/aje/kwv254

HL7. (2010). *HL7 Standards Product Brief—CDA® Release 2 | HL7 International*. http://www.hl7.org/implement/standards/product_brief.cfm?product_id=7

HL7. (2019a, November 1). *Fast Healthcare Interoperability Resources Standard*. HL7®. https://www.hl7.org/fhir/

HL7. (2019b, November 1). *Summary—FHIR v4.0.1*. https://hl7.org/FHIR/summary.html

Holshue, M. L., DeBolt, C., Lindquist, S., Lofy, K. H., Wiesman, J., Bruce, H., Spitters, C., Ericson, K., Wilkerson, S., Tural, A., Diaz, G., Cohn, A., Fox, L., Patel, A., Gerber, S. I., Kim, L., Tong, S., Lu, X., Lindstrom, S., … Invest, W. S. C. (2020). First Case of 2019 Novel Coronavirus in the United States. In *NEW ENGLAND JOURNAL OF MEDICINE* (Vol. 382, Issue 10, pp. 929–936). MASSACHUSETTS MEDICAL SOC. https://doi.org/10.1056/NEJMoa2001191

Homeland Security Council. (2006, May). *Pandemic-influenza-implementation.pdf*. https://www.cdc.gov/flu/pandemic-resources/pdf/pandemic-influenza-implementation.pdf

Hu, T., Guan, W. W., Zhu, X., Shao, Y., Liu, L., Du, J., Liu, H., Zhou, H., Wang, J., She, B., Zhang, L., Li, Z., Wang, P., Tang, Y., Hou, R., Li, Y., Sha, D., Yang, Y., Lewis, B., … Bao, S. (2020). *Building an Open Resources Repository for COVID-19 Research* (SSRN Scholarly Paper ID 3587704). Social Science Research Network. https://doi.org/10.2139/ssrn.3587704

Hughes, J. M., Wilson, M. E., Pike, B. L., Saylors, K. E., Fair, J. N., LeBreton, M., Tamoufe, U., Djoko, C. F., Rimoin, A. W., & Wolfe, N. D. (2010). The Origin and Prevention of Pandemics. *Clinical Infectious Diseases*, *50*(12), 1636–1640. https://doi.org/10.1086/652860

IHME. (2020). *Global Health Data Exchange | GHDx*. Institute for Health Metrics and Evaluation (IHME), University of Washington. http://ghdx.healthdata.org/

International Organization for Standardization. (2015). *ISO/TS 17975:2015*. ISO. https://www.iso.org/cms/render/live/en/sites/isoorg/contents/data/standard/06/11/61186.html

International Severe Acute Respiratory and Emerging Infection Consortium, & World Health Organization. (2020a, March 24). *ISARIC_COVID-19_RAPID_CRF_24MAR20_EN.pdf*. https://media.tghn.org/medialibrary/2020/04/ISARIC_COVID-19_RAPID_CRF_24MAR20_EN.pdf

International Severe Acute Respiratory and Emerging Infection Consortium, & World Health Organization. (2020b, April 23). *ISARIC_WHO_nCoV_CORE_CRF_23APR20.pdf*. https://media.tghn.org/medialibrary/2020/05/ISARIC_WHO_nCoV_CORE_CRF_23APR20.pdf

ISARIC. (2020). *COVID-19 Clinical research resources*. International Severe Acute Respiratory and Emerging INfection Consortium - The Global Health Network. https://infograph.venngage.com/pe/Mxg90X1jTc?border=false

Italian Civil Protection Department, Morettini, M., Sbrollini, A., Marcantoni, I., & Burattini, L. (2020). COVID-19 in Italy: Dataset of the Italian Civil Protection Department. *Data in Brief*, *30*, 105526. https://doi.org/10.1016/j.dib.2020.105526

Ivanov, D. (2020). Predicting the impacts of epidemic outbreaks on global supply chains: A simulation-based analysis on the coronavirus outbreak (COVID-19/SARS-CoV-2) case. *TRANSPORTATION RESEARCH PART E-LOGISTICS AND TRANSPORTATION REVIEW*, *136*. https://doi.org/10.1016/j.tre.2020.101922

Jackson, J. K., Weiss, M. A., Schwarzenberg, A. B., & Nelson, R. M. (2020). Global Economic Effects of COVID-19. *Congressional Research Service*, 2020-05–15. https://crsreports.congress.gov/product/pdf/R/R46270

Jee, Y. (2020). WHO International Health Regulations Emergency Committee for the COVID-19 outbreak. *EPIDEMIOLOGY AND HEALTH*, *42*. https://doi.org/10.4178/epih.e2020013

JHU. (2020). *COVID19 dataset* [Data repository]. Johns Hopkins University, CSSEGISandData. https://github.com/CSSEGISandData/COVID-19 (Original work published 2020)

Jiménez, R. C., Kuzak, M., Alhamdoosh, M., Barker, M., Batut, B., Borg, M., Capella-Gutierrez, S., Chue Hong, N., Cook, M., Corpas, M., Flannery, M., Garcia, L., Gelpí, J. Ll., Gladman, S., Goble, C., González Ferreiro, M., Gonzalez-Beltran, A., Griffin, P. C., Grüning, B., … Crouch, S. (2017). Four simple recommendations to encourage best practices in research software. *F1000Research*, *6*, 876. https://doi.org/10.12688/f1000research.11407.1

Kanjala, C. (2020). *Provenance of 'after the fact' harmonised community-based demographic and HIV surveillance data from ALPHA cohorts* [Doctoral, London School of Hygiene & Tropical Medicine]. https://doi.org/Kanjala, C <http://researchonline.lshtm.ac.uk/view/creators/ecpsckan.html>; (2020) Provenance of "after the fact" harmonised community-based demographic and HIV surveillance data from ALPHA cohorts. PhD thesis, London School of Hygiene & Tropical Medicine. DOI: https://doi.org/10.17037/PUBS.04655994 <https://doi.org/10.17037/PUBS.04655994>

Karesh, W. B., Dobson, A., Lloyd-Smith, J. O., Lubroth, J., Dixon, M. A., Bennett, M., Aldrich, S., Harrington, T., Formenty, P., Loh, E. H., MacHalaba, C. C., Thomas, M. J., & Heymann, D. L. (2012). Ecology of zoonoses: Natural and unnatural histories. *The Lancet*, *380*(9857), 1936–1945. Scopus. https://doi.org/10.1016/S0140-6736(12)61678-X

Kherbst, & Eehlers. (2020, April 27). *COVID-19 Screening Form 2020.pdf*. Dropbox. https://www.dropbox.com/s/iuhe4366msdfq5c/COVID-19%20Screening%20Form%202020.pdf?dl=0

Kilpatrick, A. M., & Randolph, S. E. (2012). Drivers, dynamics, and control of emerging vector-borne zoonotic diseases. *The Lancet*, *380*(9857), 1946–1955. Scopus. https://doi.org/10.1016/S0140-6736(12)61151-9

Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H., & Lipsitch, M. (2020). Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science*, *368*(6493), 860–868. https://doi.org/10.1126/science.abb5793

Knight, G., Dharan, N., & Fox, G. (2016). Bridging the gap between evidence and policy for infectious diseases: How models can aid public health decision-making. *Int J Infect Dis.*, *42*, 17–23. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4996966/

Kurt, O. K., Zhang, J., & Pinkerton, K. E. (2016). Pulmonary health effects of air pollution. *Current Opinion in Pulmonary Medicine*, *22*(2), 138–143. PubMed. https://doi.org/10.1097/MCP.0000000000000248

Kushida, C. A., Nichols, D. A., Jadrnicek, R., Miller, R., Walsh, J. K., & Griffin, K. (2012). Strategies for de-identification and anonymization of electronic health record data for use in multicenter research studies. *Medical Care*, *50*(Suppl), S82-101. https://doi.org/10.1097/MLR.0b013e3182585355

Laine, J. E., & Robinson, O. (2019). Framing Fetal and Early Life Exposome Within Epidemiology. In S. Dagnino & A. Macherone (Eds.), *Unraveling the Exposome: A Practical View* (pp. 87–123). Springer International Publishing. https://doi.org/10.1007/978-3-319-89321-1_4

Lamprecht, A.-L., Garcia, L., Kuzak, M., Martinez, C., Arcila, R., Martin Del Pico, E., Dominguez Del Angel, V., van de Sandt, S., Ison, J., Martinez, P. A., McQuilton, P., Valencia, A., Harrow, J., Psomopoulos, F., Gelpi, J. L., Chue Hong, N., Goble, C., & Capella-Gutierrez, S. (2019). Towards FAIR principles for research software. *Data Science*, *Preprint*, 1–23. https://doi.org/10.3233/DS-190026

LSRI. (2020). *LSRI Response to COVID-19*. European Life Science Research Infrastructure. https://lifescience-ri.eu/ls-ri-response-to-covid-19.html

Luis, A. D., Hayman, D. T. S., O’Shea, T. J., Cryan, P. M., Gilbert, A. T., Pulliam, J. R. C., Mills, J. N., Timonin, M. E., Willis, C. K. R., Cunningham, A. A., Fooks, A. R., Rupprecht, C. E., Wood, J. L. N., & Webb, C. T. (2013). A comparison of bats and rodents as reservoirs of zoonotic viruses: Are bats special? *Proceedings of the Royal Society B: Biological Sciences*, *280*(1756), 20122753. https://doi.org/10.1098/rspb.2012.2753

Luke, D. A., & Stamatakis, K. A. (2012). *Systems science methods in public health: Dynamics, networks, and agents*. *33*, 357–376. https://doi.org/10.1146/annurev-publhealth-031210-101222

MacFarlane, D., & Rocha, R. (2020). Guidelines for communicating about bats to prevent persecution in the time of COVID-19. *Biological Conservation*, *248*. https://doi.org/10.1016/j.biocon.2020.108650

Mahmood, S., Hasan, K., Carras, M. C., & Labrique, A. (2020). Global Preparedness Against COVID-19: We Must Leverage the Power of Digital Health. *JMIR PUBLIC HEALTH AND SURVEILLANCE*, *6*(2), 226–232. https://doi.org/10.2196/18980

Martelletti, L., & Martelletti, P. (2020). Air Pollution and the Novel Covid-19 Disease: A Putative Disease Risk Factor. *Sn Comprehensive Clinical Medicine*, 1–5. https://doi.org/10.1007/s42399-020-00274-4

Mavragani, A. (2020). Tracking COVID-19 in Europe: Infodemiology Approach. *JMIR PUBLIC HEALTH AND SURVEILLANCE*, *6*(2), 233–245. https://doi.org/10.2196/18941

McCandless, D. (2020). *COVID-19 CoronaVirus Infographic Datapack*. Information Is Beautiful. https://informationisbeautiful.net/visualizations/covid-19-coronavirus-infographic-datapack/

McMichael, T. M., Currie, D. W., Clark, S., Pogosjans, S., Kay, M., Schwartz, N. G., Lewis, J., Baer, A., Kawakami, V., Lukoff, M. D., Ferro, J., Brostrom-Smith, C., Rea, T. D., Sayre, M. R., Riedo, F. X., Russell, D., Hiatt, B., Montgomery, P., Rao, A. K., … Duchin, J. S. (2020). Epidemiology of Covid-19 in a Long-Term Care Facility in King County, Washington. *The New England Journal of Medicine*. Scopus. https://doi.org/10.1056/NEJMoa2005412

Molyneux, D., Hallaj, Z., Keusch, G. T., McManus, D. P., Ngowi, H., Cleaveland, S., Ramos-Jimenez, P., Gotuzzo, E., Kar, K., Sanchez, A., Garba, A., Carabin, H., Bassili, A., Chaignat, C. L., Meslin, F.-X., Abushama, H. M., Willingham, A. L., & Kioy, D. (2011). Zoonoses and marginalised infectious diseases of poverty: Where do we stand? *Parasites and Vectors*, *4*(1). Scopus. https://doi.org/10.1186/1756-3305-4-106

Morgan, D., & Sargent, J. F. (2020). *Effects of COVID-19 on the Federal Research and Development Enterprise* (No. R46309; CRS Reports, p. 22). Congressional Research Service (CRS), United States Library of Congress. https://crsreports.congress.gov/product/pdf/R/R46309

Morse, S. S., Mazet, J. A. K., Woolhouse, M., Parrish, C. R., Carroll, D., Karesh, W. B., Zambrana-Torrelio, C., Lipkin, W. I., & Daszak, P. (2012). Prediction and prevention of the next pandemic zoonosis. *The Lancet*, *380*(9857), 1956–1965. Scopus. https://doi.org/10.1016/S0140-6736(12)61684-5

Munthali, G. N. C., & Xuelian, W. (2020). Covid-19 Outbreak on Malawi Perspective. *ELECTRONIC JOURNAL OF GENERAL MEDICINE*, *17*(4). https://doi.org/10.29333/ejgm/7871

NASEM. (2008). *Achieving Sustainable Global Capacity for Surveillance and Response to Emerging Diseases of Zoonotic Origin: Workshop Summary*. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/12522

NASEM. (2009). *Sustaining Global Surveillance and Response to Emerging Zoonotic Diseases*. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/12625

NASEM. (2010). *Infectious Disease Movement in a Borderless World: Workshop Summary*. National Academies of Sciences, Engineering, and Medicine. https://www.nap.edu/download/12758

NASEM. (2015). *Emerging Viral Diseases: The One Health Connection: Workshop Summary*. National Academies of Sciences, Engineering, and Medicine. https://www.nap.edu/catalog/18975/emerging-viral-diseases-the-one-health-connection-workshop-summary

NASEM. (2018). *Exploring Lessons Learned from Partnerships to Improve Global Health and Safety: Workshop in Brief*. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/21690

NASEM. (2020a). *Coronavirus resources collection*. National Academies of Sciences, Engineering, and Medicine. http://www.nap.edu/collection/94/coronavirus-resources

NASEM. (2020b). *Rapid Expert Consultation on Data Elements and Systems Design for Modeling and Decision Making for the COVID-19 Pandemic (March 21, 2020)*. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25755

NASEM. (2020c). *Rapid Expert Consultations on the COVID-19 Pandemic: March 14, 2020-April 8, 2020* (National Academies of Sciences, Engineering, and Medicine). National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25784

NASEM. (2020d). *Evaluating Data Types: A Guide for Decision Makers using Data to Understand the Extent and Spread of COVID-19*. National Academies of Sciences, Engineering, and Medicine. https://doi.org/10.17226/25826

NASEM, Keusch, G. T., Pappaioanou, M., Gonzalez, M. C., Scott, K. A., & Tsai, P. (2009). Achieving an Effective Zoonotic Disease Surveillance System. In *Sustaining Global Surveillance and Response to Emerging Zoonotic Diseases*. National Academies of Sciences, Engineering, and Medicine. https://www.ncbi.nlm.nih.gov/books/NBK215315/

National Institute of Allergy and Infectious Disease (NIAID). (2013). *Data Sharing and Release Guidelines*. https://www.niaid.nih.gov/research/data-sharing-and-release-guidelines

Nelson, C., Lurie, N., Wasserman, J., & Zakowski, S. (2007). Conceptualizing and defining public health emergency preparedness. *American Journal of Public Health*, *97 Suppl 1*, S9-11. Scopus. https://doi.org/10.2105/AJPH.2007.114496

Ng, V., & Sargeant, J. M. (2013). A Quantitative Approach to the Prioritization of Zoonotic Diseases in North America: A Health Professionals' Perspective. In *PLOS ONE* (Vol. 8, Issue 8). PUBLIC LIBRARY SCIENCE. https://doi.org/10.1371/journal.pone.0072172

NIH. (2020a). *ClinicalTrials—Listed clinical studies related to the coronavirus disease (COVID-19)*. U.S. National Institutes of Health - Information on Clinical Trials and Human Research Studies - National Library of Medicine. https://clinicaltrials.gov/ct2/results?cond=COVID-19

NIH. (2020b). *NIH Public Health Emergency and Disaster Research Response (DR2) COVID-19 Research Tools—Training Material*. NIH Public Health Emergency and Disaster Research Response (DR2). https://dr2.nlm.nih.gov/

NIH. (2020c). *COVID-19 OBSSR Research Tools*. National Institutes of Health. https://www.nlm.nih.gov/dr2/COVID-19_BSSR_Research_Tools.pdf

Nii-Trebi, N. I. (2017). Emerging and Neglected Infectious Diseases: Insights, Advances, and Challenges. *BioMed Research International*, *2017*. https://doi.org/10.1155/2017/5245021

NIST PWG. (2015). *Big Data Interoperability Framework: Volume 5, Architectures White Paper Survey* (NIST SP 1500-5; p. NIST SP 1500-5, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-5

NIST PWG. (2019a). *Big Data Interoperability Framework: Volume 1, Definitions* (NIST SP 1500-1r2; p. NIST SP 1500-1r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-1r2

NIST PWG. (2019b). *Big Data Interoperability Framework: Volume 3, Use Cases and General Requirements* (NIST SP 1500-3r2; p. NIST SP 1500-3r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-3r2

NIST PWG. (2019c). *Big Data Interoperability Framework: Volume 4, Security and Privacy* (NIST SP 1500-4r2; p. NIST SP 1500-4r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-4r2

NIST PWG. (2019d). *Big Data Interoperability Framework: Volume 6, Reference Architecture* (NIST SP 1500-6r2; p. NIST SP 1500-6r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-6r2

NIST PWG. (2019e). *Big Data Interoperability Framework: Volume 7, Standards Roadmap* (NIST SP 1500-7r2; p. NIST SP 1500-7r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-7r2

NIST PWG. (2019f). *Big Data Interoperability Framework: Volume 8, Reference Architecture Interfaces* (NIST SP 1500-9r1; p. NIST SP 1500-9r1, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-9r1

NIST PWG. (2019g). *Big Data Interoperability Framework: Volume 9, Adoption and Modernization* (NIST SP 1500-10r1; p. NIST SP 1500-10r1, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-10r1

NIST PWG. (2019h). *Big Data Interoperability Framework: Volume 2, Big Data Taxonomies* (NIST SP 1500-2r2; p. NIST SP 1500-2r2, version 3 (final)). National Institute of Standards and Technology, Big Data Public Working Group. https://doi.org/10.6028/NIST.SP.1500-2r2

NSW Health. (2020, March 23). *Novel-coronavirus-case-questionnaire.pdf*. https://www.health.nsw.gov.au/Infectious/Forms/novel-coronavirus-case-questionnaire.pdf
NYC Health. (2020). *Nychealth/coronavirus-data*. NYC Department of Health and Mental Hygiene. https://github.com/nychealth/coronavirus-data (Original work published 2020)
NYT. (2020, April 21). *Coronavirus (Covid-19) Data in the United States*. New York Times. https://github.com/nytimes/covid-19-data (Original work published 2020)
OECD. (2019). *Recommendation of the Council on Health Data Governance*. https://www.oecd.org/health/health-systems/Recommendation-of-OECD-Council-on-Health-Data-Governance-Booklet.pdf
Ogunyemi, O. I., Meeker, D., Kim, H.-E., Ashish, N., Farzaneh, S., & Boxwala, A. (2013). Identifying Appropriate Reference Data Models for Comparative Effectiveness Research (CER) Studies Based on Data from Clinical Information Systems. *MEDICAL CARE*, *51*(8, 3), S45–S52. https://doi.org/10.1097/MLR.0b013e31829b1e0b
OHDSI. (2019). *OMOP Common Data Model – OHDSI*. Observational Health Data Sciences and Informatics. https://www.ohdsi.org/data-standardization/the-common-data-model/
OpenAIRE. (2020). *COVID-19 Open Research Gateway*. OpenAIRE - Connect. https://beta.covid-19.openaire.eu/content
Oxford University. (2020a). *COVID19 dataset* [Dataset]. https://github.com/owid/covid-19-data
Oxford University. (2020b). *COVID19 government response tracker* [Dataset]. University of Oxford. https://www.bsg.ox.ac.uk/research/research-projects/coronavirus-government-response-tracker
Panchadsaram, R., Davis, N., Wikelius, K., Shahpar, C., Cobb, L., Wosińska, M., Anderson, D., Brown, L. M., Hunt, D., & Goligorsky, D. (2020). *COVID exit strategy: How we reopen safely*. https://www.covidexitstrategy.org/
Park, H. W., Park, S., & Chong, M. (2020). Conversations and Medical News Frames on Twitter: Infodemiological Study on COVID-19 in South Korea. *JOURNAL OF MEDICAL INTERNET RESEARCH*, *22*(5). https://doi.org/10.2196/18897
Pathak, E. B., Salemi, J. L., Sobers, N., Menard, J., & Hambleton, I. R. (2020). COVID-19 in Children in the United States: Intensive Care Admissions, Estimated Total Infected, and Projected Numbers of Severe Pediatric Cases in 2020. *Journal of Public Health Management and Practice*, *Publish Ahead of Print*. https://doi.org/10.1097/PHH.0000000000001190
PCORnet. (2020). *Patient-Centered Outcomes Research Institute*. The National Patient-Centered Clinical Research Network. https://pcornet.org/
Peirlinck, M., Linka, K., Sahli Costabal, F., & Kuhl, E. (2020). Outbreak dynamics of COVID-19 in China and the United States. *BIOMECHANICS AND MODELING IN MECHANOBIOLOGY*. https://doi.org/10.1007/s10237-020-01332-5
periCOVID Uganda CRF. (2020, April 10). *PeriCOVID Uganda CRF.docx*. Dropbox. https://www.dropbox.com/s/1p32oudodv8bm1h/periCOVID%20Uganda%20CRF.docx?dl=0
PhenX. (2020). *PhenX Toolkit—COVID-19 Protocols*. https://www.phenxtoolkit.org/covid19
Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., & Lloyd-Smith, J. O. (2017). Pathways to zoonotic spillover. *Nature Reviews Microbiology*, *15*(8), 502–510. https://doi.org/10.1038/nrmicro.2017.45
Priyadarsini, S. L., & Suresh, M. (2020). Factors influencing the epidemiological characteristics of pandemic COVID 19: A TISM approach. *INTERNATIONAL JOURNAL OF HEALTHCARE MANAGEMENT*. https://doi.org/10.1080/20479700.2020.1755804

Ray, D., Salvatore, M., Bhattacharyya, R., Wang, L., Du, J., Mohammed, S., Purkayastha, S., Halder, A., Rix, A., Barker, D., Kleinsasser, M., Zhou, Y., Bose, D., Song, P., Banerjee, M., Baladandayuthapani, V., Ghosh, P., & Mukherjee, B. (2020). Predictions, Role of Interventions and Effects of a Historic National Lockdown in India's Response to the the COVID-19 Pandemic: Data Science Call to Arms. *Harvard Data Science Review*. https://doi.org/10.1162/99608f92.60e08ed5

RDA-COVID19 Zotero WG. (2020, June 5). *RDA-COVID19 WG Zotero Library*. RDA. https://doi.org/10.15497/rda00051

RDA-COVID19-Epidemiology WG. (2020). *Data sharing in epidemiology (version 0.06b)*. https://www.rd-alliance.org/group/rda-covid19-epidemiology/outcomes/data-sharing-epidemiology

Ricciardi, F., De Bernardi, P., & Cantino, V. (2020). System dynamics modeling as a circular process: The smart commons approach to impact management. *Technological Forecasting and Social Change*, *151*(C). https://ideas.repec.org/a/eee/tefoso/v151y2020ics0040162519310923.html

Rist, C. L., Arriola, C. S., & Rubin, C. (2014). Prioritizing Zoonoses: A Proposed One Health Tool for Collaborative Decision-Making. In *PLOS ONE* (Vol. 9, Issue 10). PUBLIC LIBRARY SCIENCE. https://doi.org/10.1371/journal.pone.0109986

Rockefeller Foundation. (2020, May 19). *Call for Entries: Data Science Breakthroughs for an Inclusive Recovery*. The Rockefeller Foundation. https://www.rockefellerfoundation.org/blog/call-for-entries-data-science-breakthroughs-for-an-inclusive-recovery/

Sansone, S.-A., Rocca-Serra, P., Field, D., Maguire, E., Taylor, C., Hofmann, O., Fang, H., Neumann, S., Tong, W., Amaral-Zettler, L., Begley, K., Booth, T., Bougueleret, L., Burns, G., Chapman, B., Clark, T., Coleman, L.-A., Copeland, J., Das, S., … Hide, W. (2012). Toward interoperable bioscience data. *Nature Genetics*, *44*(2), 121–126. https://doi.org/10.1038/ng.1054

Saulnier, K. M., Bujold, D., Dyke, S. O. M., Dupras, C., Beck, S., Bourque, G., & Joly, Y. (2019). Benefits and barriers in the design of harmonized access agreements for international data sharing. *Scientific Data*, *6*(1), 297. https://doi.org/10.1038/s41597-019-0310-4

Semantic Scholar. (2020). *CORD-19*. https://pages.semanticscholar.org/coronavirus-research

Setti, L., Passarini, F., Gennaro, G. D., Baribieri, P., Perrone, M. G., Borelli, M., Palmisani, J., Gilio, A. D., Torboli, V., Pallavicini, A., Ruscio, M., Piscitelli, P., & Miani, A. (2020). *SARS-Cov-2 RNA Found on Particulate Matter of Bergamo in Northern Italy: First Preliminary Evidence*. Cold Spring Harbor Laboratory Press. https://www.medrxiv.org/content/10.1101/2020.04.15.20065995v2

Shiferaw, M. L., Doty, J. B., Maghlakelidze, G., Morgan, J., Khmaladze, E., Parkadze, O., Donduashvili, M., Wemakoy, E. O., Muyembe, J.-J., Mulumba, L., Malekani, J., Kabamba, J., Kanter, T., Boulanger, L. L., Haile, A., Bekele, A., Bekele, M., Tafese, K., McCollum, A. M., & Reynolds, M. G. (2017). *Frameworks for Preventing, Detecting, and Controlling Zoonotic Diseases—Volume 23, Supplement—December 2017—Emerging Infectious Diseases journal—CDC*. https://doi.org/10.3201/eid2313.170601

Smiley Evans, T., Shi, Z., Boots, M., Liu, W., Olival, K. J., Xiao, X., Vandewoude, S., Brown, H., Chen, J.-L., Civitello, D. J., Escobar, L., Grohn, Y., Li, H., Lips, K., Liu, Q., Lu, J., Martínez-López, B., Shi, J., Shi, X., … Getz, W. M. (2020). Synergistic China–US Ecological Research is Essential for Global Emerging Infectious Disease Preparedness. *EcoHealth*, *17*(1), 160–173. Scopus. https://doi.org/10.1007/s10393-020-01471-2

Smith, K. F., Behrens, M., Schloegel, L. M., Marano, N., Burgiel, S., & Daszak, P. (2009). Reducing the risks of the wildlife trade. *Science*, *324*(5927), 594–595. Scopus. https://doi.org/10.1126/science.1174460

SNOMED. (2020). *COVID-19 Data Coding using SNOMED CT - COVID-19 Guide—SNOMED Confluence*. https://confluence.ihtsdotools.org/display/DOCCV19

Staunton, C., Slokenberga, S., & Mascalzoni, D. (2019). The GDPR and the research exemption: Considerations on the necessary safeguards for research biobanks. *European Journal of Human Genetics*, *27*(8), 1159–1167. https://doi.org/10.1038/s41431-019-0386-5

Sun, P., Lu, X., Xu, C., Sun, W., & Pan, B. (2020). Understanding of COVID-19 based on current evidence. *Journal of Medical Virology*, *92*(6), 548–551. https://doi.org/10.1002/jmv.25722

Taylor, L. H., Latham, S. M., & Woolhouse, M. E. J. (2001). Risk factors for human disease emergence. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *356*(1411), 983–989. https://doi.org/10.1098/rstb.2001.0888

Team, G. P. (2020). *GESIS Panel Special Survey on the Coronavirus SARS-CoV-2 Outbreak in Germany*. https://doi.org/10.4232/1.13520

Templ, M. (2015). Quality indicators for statistical disclosure methods: A case study on the structure of earnings survey. *Journal of Official Statistics*, *31*(4), 737–761. Scopus. https://doi.org/10.1515/JOS-2015-0043

Templ, M. (2017). *Statistical disclosure control for microdata: Methods and applications in R* (p. 287). Springer International Publishing; Scopus. https://doi.org/10.1007/978-3-319-50272-4

Templ, Matthias, Meindl, B., & Kowarik, A. (2020). *sdcMicro: Statistical Disclosure Control Methods for Anonymization of Data and Risk Estimation*. https://cran.r-project.org/web/packages/sdcMicro/index.html

Templ, Matthias, Meindl, B., Kowarik, A., & Chen, S. (2014). *Introduction to Statistical Disclosure Control (SDC)*. 25. https://www.ihsn.org/sites/default/files/resources/ihsn-working-paper-007-Oct27.pdf

The Atlantic. (2020a). *COVID Tracking Project* [Datasets]. https://covidtracking.com/

The Atlantic. (2020b). *The COVID Racial Data Tracker*. The COVID Tracking Project. https://covidtracking.com/race

The European Data Protection Board. (2020). *Guidelines in the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak*. https://edpb.europa.eu/sites/edpb/files/files/file1/edpb_guidelines_202003_healthdatascientificresearchcovid19_en.pdf

The European Health Data & Evidence Network. (2020, April 15). *COVID-19-Data-Partner-Call-Description-v1.4.pdf*. https://www.ehden.eu/wp-content/uploads/2020/04/COVID-19-Data-Partner-Call-Description-v1.4.pdf

The National Institute of Environmental Health Science. (2020, May 14). *COVID-19_BSSR_Research_Tools.pdf*. https://www.nlm.nih.gov/dr2/COVID-19_BSSR_Research_Tools.pdf

UN. (2018). *The Humanitarian Exchange Language (HXL)*. United Nations, Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. https://hxlstandard.org/standard/1-1final/

UN. (2020). *The Humanitarian Data Exchange (HDX)*. United Nations, Office for the Coordination of Humanitarian Affairs (OCHA), Centre for Humanitarian Data. https://data.humdata.org/

UNDRR. (2020). *Disaster risk management for health: Overview*. https://www.undrr.org/publication/disaster-risk-management-health-overview

Universidade Federal de Pelotas. (2020). *Brazil COVID serological survey questionnaire 20200428.docx*. Dropbox.

https://www.dropbox.com/s/l9hhblb83ybr70u/Brazil%20COVID%20serological%20survey%20questionnaire%2020200428.docx?dl=0
University of Maryland. (2020, April 24). *COVID-19 Impact Analysis Platform*. COVID-19 Impact Analysis Platform. https://data.covid.umd.edu/
University of Washington. (2020). *COVID19 data: Beoutbreakprepared* [Data repository]. https://github.com/beoutbreakprepared
University of Washington - HGIS Lab. (2020). *Novel Coronavirus Infection Map*. University of Washingtion - Humanistic GIS Laboratory. https://github.com/jakobzhao/virus
U.S. Department of Health and Human Services. (2017, December). *Pan-flu-report-2017v2.pdf*. https://www.cdc.gov/flu/pandemic-resources/pdf/pan-flu-report-2017v2.pdf
van Bochove, K., Vos, E., van Winzum, A., Kurps, J., & Moinat, M. (2020). *Implementing FAIR in OHDSI: Challenges and opportunities for EHDEN*. 1.
Voss, E. A., Makadia, R., Matcho, A., Ma, Q., Knoll, C., Schuemie, M., DeFalco, F. J., Londhe, A., Zhu, V., & Ryan, P. B. (2015). Feasibility and utility of applications of the common data model to multiple, disparate observational health databases. *JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION*, *22*(3), 553–564. https://doi.org/10.1093/jamia/ocu023
Wang, L. L., Lo, K., Chandrasekhar, Y., Reas, R., Yang, J., Eide, D., Funk, K., Kinney, R., Liu, Z., Merrill, W., Mooney, P., Murdick, D., Rishi, D., Sheehan, J., Shen, Z., Stilson, B., Wade, A. D., Wang, K., Wilhelm, C., … Kohlmeier, S. (2020). CORD-19: The Covid-19 Open Research Dataset. *ArXiv:2004.10706 [Cs]*. http://arxiv.org/abs/2004.10706
Warren, M. (2020, April 20). *CDISC Interim User Guide for COVID-19—CDISC Interim User Guide for COVID-19—Wiki*. https://wiki.cdisc.org/display/COVID19/CDISC+Interim+User+Guide+for+COVID-19
Webster, R. G. (2004). Wet markets—A continuing source of severe acute respiratory syndrome and influenza? *The Lancet*, *363*(9404), 234–236. https://doi.org/10.1016/S0140-6736(03)15329-9
Wellcome. (2020, April 23). *Final UK Covid Questionnaire_23 April.pdf*. Dropbox. https://www.dropbox.com/s/hy6jdpgvkgpgxfi/Final%20UK%20Covid%20Questionnaire_23%20April.pdf?dl=0
Wellcome Trust. (2017). *Longitudinal Population Studies Strategy*. https://wellcome.ac.uk/sites/default/files/longitudinal-population-studies-strategy_0.pdf
WHO. (2003). *Climate change and infectious diseases*. Climate Change and Human Health; World Health Organization. https://www.who.int/globalchange/summary/en/index5.html
WHO. (2006). *Global Early Warning System for Major Animal Diseases, including Zoonoses (GLEWS)*. World Health Organization. https://www.who.int/foodsafety/areas_work/zoonose/glews/en/
WHO. (2015, September 2). *Developing Global Norms for Sharing Data and Results during Public Health Emergencies*. World Health Organization; World Health Organization. http://www.who.int/medicines/ebola-treatment/data-sharing_phe/en/
WHO. (2017a). *One Health*. https://www.who.int/news-room/q-a-detail/one-health
WHO. (2017b, July 19). *Zoonoses*. WHO; World Health Organization. http://www.who.int/topics/zoonoses/en/
WHO. (2020a). *About EPI-WIN*. https://www.who.int/teams/risk-communication/about-epi-win
WHO. (2020b). *COVID-19 situation reports*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

WHO. (2020c). *Ethical considerations to guide the use of digital proximity tracking technologies for COVID-19 contact tracing*. https://www.who.int/publications-detail/WHO-2019-nCoV-Ethics_Contact_tracing_apps-2020.1WHO

WHO. (2020d). *Novel Coronavirus (2019-nCoV) situation reports*. World Health Organization. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

WHO. (2020e). *Population-based age-stratified seroepidemiological investigation protocol for COVID-19 virus infection*. https://apps.who.int/iris/bitstream/handle/10665/332188/WHO-2019-nCoV-Seroepidemiology-2020.2-eng.pdf?sequence=1&isAllowed=y

WHO. (2020f). *Risk Communication: EPI-WIN*. https://www.who.int/teams/team-preview/risk-communciation---global

WHO. (2020g). *Survey tool and guidance: Behavioural insights on COVID-19*. WHO Regional Office for Europe. http://www.euro.who.int/__data/assets/pdf_file/0007/436705/COVID-19-survey-tool-and-guidance.pdf?ua=1

WHO. (2020h). *WHO | Global Influenza Surveillance and Response System (GISRS)*. WHO; World Health Organization. http://www.who.int/influenza/gisrs_laboratory/en/

WHO. (2020i, February). *COVID-19 CRF • ISARIC*. https://isaric.tghn.org/COVID-19-CRF/

WHO. (2020j). *Report of the WHO-China Joint Mission on Coronavirus Disease 2019 (COVID-19)*. World Health Organization. https://www.who.int/publications-detail/report-of-the-who-china-joint-mission-on-coronavirus-disease-2019-(covid-19)

WHO. (2020k, March 20). *Global surveillance for COVID-19 caused by human infection with COVID-19 virus: Interim guidance*. https://apps.who.int/iris/bitstream/handle/10665/331506/WHO-2019-nCoV-SurveillanceGuidance-2020.6-eng.pdf

WHO. (2020l, March 20). *Surveillance, rapid response teams, and case investigation*. Coronavirus Disease (COVID-19) Technical Guidance. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/surveillance-and-case-definitions

WHO. (2020m, March 23). *WHO Global COVID-19 Clinical Platform Case Record Form (CRF)*. https://www.who.int/publications-detail/global-covid-19-clinical-platform-novel-coronavius-(-covid-19)-rapid-version

WHO. (2020n, April 29). *Modes of transmission of virus causing COVID-19: Implications for IPC precaution recommendations*. https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations

WHO. (2020o, May 26). *Preparing GISRS for the upcoming influenza seasons during the COVID-19 pandemic – practical considerations*. https://apps.who.int/iris/bitstream/handle/10665/332198/WHO-2019-nCoV-Preparing_GISRS-2020.1-eng.pdf

WHO. (2020p, May 29). *C-TAP: COVID-19 technology access pool*. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/global-research-on-novel-coronavirus-2019-ncov/covid-19-technology-access-pool

WHO, FAO and OIE. (2019). *Taking a Multisectoral, One Health Approach: A Tripartite Guide to Addressing Zoonotic Diseases in Countries*. https://www.oie.int/fileadmin/Home/eng/Media_Center/docs/EN_TripartiteZoonosesGuide_webversion.pdf

Wicher, D. (2020). *The covid-19 case as an example of Systems Thinking usage*. https://agilejar.com/2020/03/the-covid-19-case-as-an-example-of-systems-thinking-usage/

Woo, P. C., Lau, S. K., & Yuen, K. (2006). Infectious diseases emerging from Chinese wet-markets: Zoonotic origins of severe respiratory viral infections. *Current Opinion in Infectious Diseases*, *19*(5), 401–407. https://doi.org/10.1097/01.qco.0000244043.08264.fc

World Bank. (2020). *Understanding the Coronavirus (COVID-19) pandemic through data* [Datasets]. http://datatopics.worldbank.org/universal-health-coverage/covid19/

Worldometer. (2020). *COVID19 data* [Dataset]. https://www.worldometers.info/coronavirus/

Wu, D., Wu, T., Liu, Q., & Yang, Z. (2020). The SARS-CoV-2 outbreak: What we know. *International Journal of Infectious Diseases*, *94*, 44–48. Scopus. https://doi.org/10.1016/j.ijid.2020.03.004

Xu, B., Gutierrez, B., Mekaru, S., Sewalk, K., Goodwin, L., Loskill, A., Cohn, E. L., Hswen, Y., Hill, S. C., Cobo, M. M., Zarebski, A. E., Li, S., Wu, C.-H., Hulland, E., Morgan, J. D., Wang, L., O'Brien, K., Scarpino, S. V., Brownstein, J. S., … Kraemer, M. U. G. (2020). Epidemiological data from the COVID-19 outbreak, real-time case information. *Scientific Data*, *7*(1), 106. https://doi.org/10.1038/s41597-020-0448-0

Yang, C., Qiu, X., Fan, H., Jiang, M., Lao, X., Zeng, Y., & Zhang, Z. (2020). Coronavirus disease 2019: Reassembly attack of coronavirus. *INTERNATIONAL JOURNAL OF ENVIRONMENTAL HEALTH RESEARCH*. https://doi.org/10.1080/09603123.2020.1747602

Yang, T., Shen, K., He, S., Li, E., Sun, P., Chen, P., Zuo, L., Hu, J., Mo, Y., Zang, W., Zhang, H., Chen, J.-X., & Guo, Y. (2020). *CovidNet: 1Point3Acres*. https://coronavirus.1point3acres.com/en

Yang, T., Shen, K., He, S., Li, E., Sun, P., Chen, P., Zuo, L., Hu, J., Mo, Y., Zhang, W., Zhang, H., Chen, J., & Guo, Y. (2020). *CovidNet: To bring data transparency in the era of COVID-19*. http://arxiv.org/abs/2005.10948

Yasaka, T. M., Lehrich, B. M., & Sahyouni, R. (2020). Peer-to-Peer Contact Tracing: Development of a Privacy-Preserving Smartphone App. *JMIR MHEALTH AND UHEALTH*, *8*(4). https://doi.org/10.2196/18936

Zastrow, M. (2020). Open science takes on the coronavirus pandemic. *Nature*, *581*(7806), 109–110. https://doi.org/10.1038/d41586-020-01246-3

Zhang, L., Ghader, S., Pack, M. L., Xiong, C., Darzi, A., Yang, M., Sun, Q., Kabiri, A., & Hu, S. (2020). An interactive COVID-19 mobility impact and social distancing analysis platform. *MedRxiv*, 2020.04.29.20085472. https://doi.org/10.1101/2020.04.29.20085472

Zhang, Y. (2020). *The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19)—China, 2020*. Cina CDC Weekly. http://weekly.chinacdc.cn/en/article/id/e53946e2-c6c4-41e9-9a9b-fea8db1a8f51

Zhou, P., Yang, X.-L., Wang, X.-G., Hu, B., Zhang, L., Zhang, W., Si, H.-R., Zhu, Y., Li, B., Huang, C.-L., Chen, H.-D., Chen, J., Luo, Y., Guo, H., Jiang, R.-D., Liu, M.-Q., Chen, Y., Shen, X.-R., Wang, X., … Shi, Z.-L. (2020). A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*, *579*(7798), 270–273. https://doi.org/10.1038/s41586-020-2012-7

Zhou, S.-J., Zhang, L.-G., Wang, L.-L., Guo, Z.-C., Wang, J.-Q., Chen, J.-C., Liu, M., Chen, X., & Chen, J.-X. (2020). Prevalence and socio-demographic correlates of psychological health problems in Chinese adolescents during the outbreak of COVID-19. *EUROPEAN CHILD & ADOLESCENT PSYCHIATRY*. https://doi.org/10.1007/s00787-020-01541-4

Zhou, Y., Wang, L., Zhang, L., Shi, L., Yang, K., He, J., Zhao, B., Overton, W., Purkayastha, S., Song, P., Song, P., & Purkayastha, S. (2020). A Spatiotemporal Epidemiological Prediction Model to Inform County-level COVID-19 Risk in the USA. *Harvard Data Science Review*. https://hdsr.mitpress.mit.edu/pub/qqg19a0r/release/1