Week 3: Data Collection Methods in COVID Research

Nate Bachmeier

TIM-7211: Research Design Methodologies

January 3rd, 2021

Northcentral University

# Data Collection Methods in COVID Research

The novel coronavirus pandemic has infected nearly 85 million people globally and severely impacted countless businesses (Johns Hopkins, 2020). One of the challenges that come with a *novel* virus is the lack of available information. Researchers around the world stepped forward to fill these gaps through a collection of different data collection methods. During the initial outbreak, information was scarce, making qualitative approaches more appropriate (Valadez et al., 2020). Others published quantitative journals using generic biology models (e.g., Hill functions) or aggregated data from multiple sources. Several of these naïve models were not sufficient in hindsight, resulting in multiple revisions to the covid guidelines (Center for Disease Control and Prevention, 2020) and a confused population. However, this incident can serve as a useful example for data collection methods in highly ambiguous scenarios.

## Preparedness of Our Emergency Department

Hou et al. (November 2020) conducted a postmortem review on the nurses’ perspective of emergency room change adoption. The researchers wanted to understand specifically (1) how did hospitals operate at the initial onset, (2) what changes became necessary, and (3) did those changes place promptly. Their team began by documenting an open-ended interview and a series of follow-up probing questions. Then, they transcribed verbatim one-on-one sessions with dozens of nurses across three hospitals in the same region. Afterward, they assigned two researchers to extract themes from the transcriptions independently. Next, cross-checking the results ensures the extractions were consistent. Finally, a mapping of critical ideas into a hierarchical structure occurred.

Their qualitative assessment shows that during the onset, hospitals were not very efficient. However, frequent updates from leadership instilled confidence to adopt workflow changes quickly. Several nurses experienced distress and isolation because they did not want to concern friends and family. Another observation is that multidisciplinary collaboration reached an all-time high, and everyone felt they were in this together.

## Modeling the Outbreak

In March and April 2020, governments across the globe began to enforce mass quarantine procedures. This action disturbed schools and businesses, leading to the unanimous question on everyone’s mind, “when will this end?” However, an authoritative dataset did not exist nor even consistent reporting of infection rates.

Jiang et al. (May 2020) understood four stages of the epidemic curve: where a virus begins sporadically, infects locally, emerges broadly, and then flattens into irrelevance. They used daily-infection rates per country to assess the stage of each nation. Next, the researchers calculated each nation’s progress percentage as (a) daily total confirmed cases by (b) up-to-date total cases. This metric’s accuracy was confirmed by measuring durations of stage-3 and stage-4 nations.

However, their model predicts, “the epidemic would end before August 10th, 2020 in the USA […] with total cases not exceeding 2.50 million (Jiang et al., 2020, p. 1109).” This purely quantitative method was inaccurate due to ignoring many external factors. For instance, the massive quarantine reshaped the curve, and they assumed that all countries would respond homogenously. Instead, using a mixed-method data collection could have identified these risks.

## Progress and Challenges with Drugs

El-Aziz & Stockand (April 2020) wrote a narrative that follows the infection spreading from bats to snakes, to humans, and finally, other humans. At each stage of the supply chain, they describe the biological processes involved then dive into a quantitative assessment of any relevant systems. For example, a person might become ill after contacting a contaminated surface. Different surface types (e.g., cardboard versus steal) can support the virus for various durations. The authors address these distinctions one-by-one by aggregating numerous academic publications that cross-validate each other.

After assessing how the virus spreads, the researchers drill into the epidemiology to determine risk rates to various cohorts (e.g., by age or gender). They accomplish this task by collecting global daily infection rates. However, it is too early into the pandemic for researchers to confirm that the preliminary figures are accurate. The authors compute various metrics on these counts, then evaluate the results against SARS and MERS historical data. Finally, they conclude that “more structural biology details about the 2019-nCoV life cycle are needed (El-Aziz & Stockand, 2020, p. 9) […] to accelerate the development of a vaccine.”

## Modeling Cell Interaction

When a person falls sick, quarantining the patient prevents spreading the disease further, but how long must they be isolated? Sohail & Nutini (March 2020) met this challenge by reviewing the sequence of events from both the SARS (Severe Acute Respiratory Syndrome) epidemic of 2002 and the MERS (Middle East Respiratory Syndrome) of 2012. Both incidents were consistent with *Hill functions*, a generic statistical model for biomedical infection rates. The two most essential parameters into these Hill functions are (1) the infectious dose (e.g., exposure rate) and (2) the incubation period (e.g., the delay before the illness). Unfortunately, these values were unknown, so the researchers devised a Markov Chain based Monte Carlo Simulation to approximate the ranges. These simulations recommended that quarantining sick individuals for roughly twelve days was the ideal period. Today, after collecting millions of real-world examples, doctors continue to give similar guidance.

## Effect on Hospitality Industry

## Effect of Confinement