Public-Private Partnerships for Data-Informed Decision Making to Distribute, Dispense, and Administer Medical Countermeasures

NASEM Public Workshop | October 3, 2023

Jessica Malaty Rivera, MS
Senior Science Communication Advisor, The deBeaumont Foundation
Research Assistant and Doctoral Student, Johns Hopkins University Center for Health Security

The COVID Tracking Project

From 2020 to 2021, the COVID Tracking Project compiled and published the most complete data about COVID-19 in the US.

Our volunteer teams analyzed testing, hospitalization, and outcomes data from 56 official US state and territorial jurisdictions.

The COVID Tracking Project

In the process of compiling daily COVID-19 data, we developed several research habits that helped us interpret the data we collected in a responsible, cautious way.

The COVID Tracking Project

We did all the updates by hand, which allowed us to maintain close touch with annotations, definition changes, anomalies, reliable and artificial trends, caveats, and disclaimers.

- Testing & outcomes: daily
- Race & Ethnicity: 2x a week
- Long Term Care: 1x a week

Data Reporting Priorities

The COVID Tracking Project

- **Inclusivity:** full access, regardless of ability, device, or connection speed
- **Performance:** all pages lean and performant, with low-end devices, mobile use, and shaky connections in mind
- **Transparency:** Surface information easily—no hiding data behind hover effects, and make sure anything offered as a map is also provided as a table and as a downloadable CSV.

The COVID Tracking Project

Data pipelines at the local, state, and national levels had serious *and different* flaws. State data was often more comprehensive but less comparable than federal data.

The COVID Tracking Project

Throughout the year, officials and members of the public made decisions based on incomplete, unstandardized, and sometimes inaccurate data.

The COVID Tracking Project

Even where data existed, federal agencies failed to communicate it in clear, useful, or even accurate ways.

This created honest confusion and opportunities for bad-faith actors to spread misinformation.

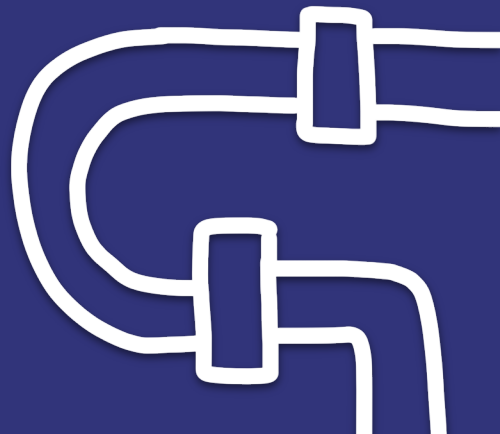
Data Reporting Best Practices

The COVID Tracking Project

- **Prioritize clarity:** Putting the most important data point associated with each metric front and center

- **Building trust through transparency:** Offering obvious, accessible ways for visitors to learn more about each metric's meaning and sources. Offer a historical time-series for all metrics and demographic data, and let people easily download and explore it

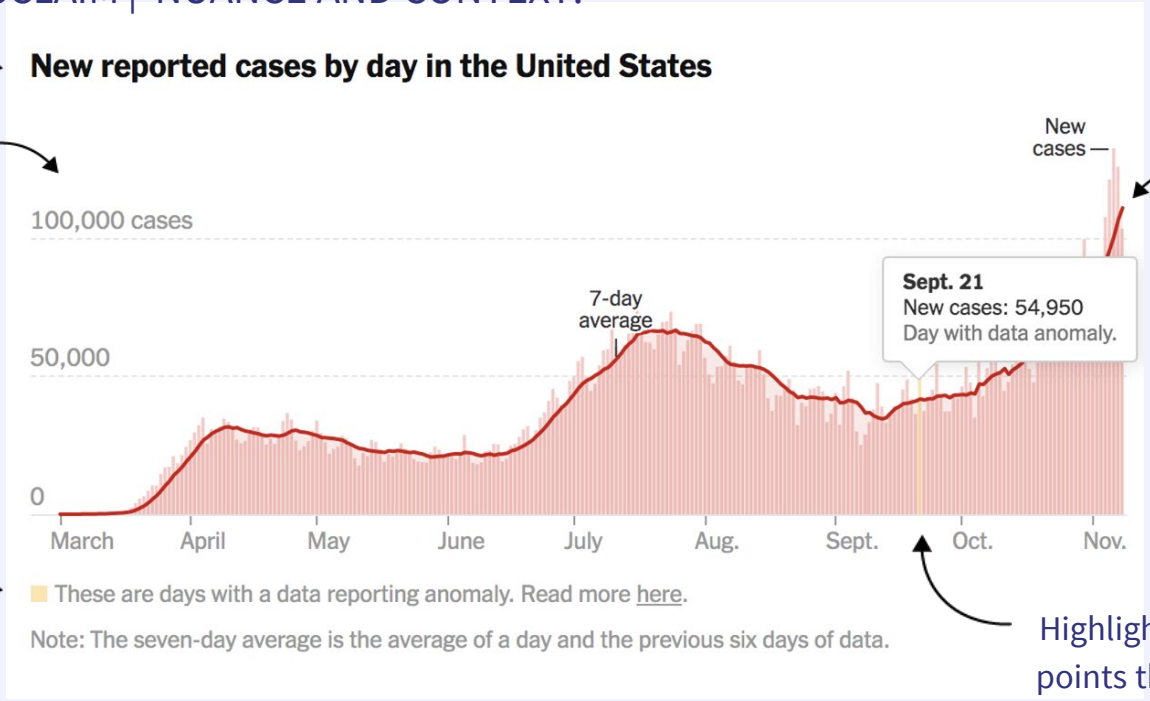
- **Using structure and consistency to boost understanding:** Thoughtfully grouping metrics and consistently report data points in the same place and in the same way, every time



Graphics that Build Understanding and Trust

ANNOTATE AND DISCLAIM | NUANCE AND CONTEXT.

Use clear titles, legends, and labels for charts

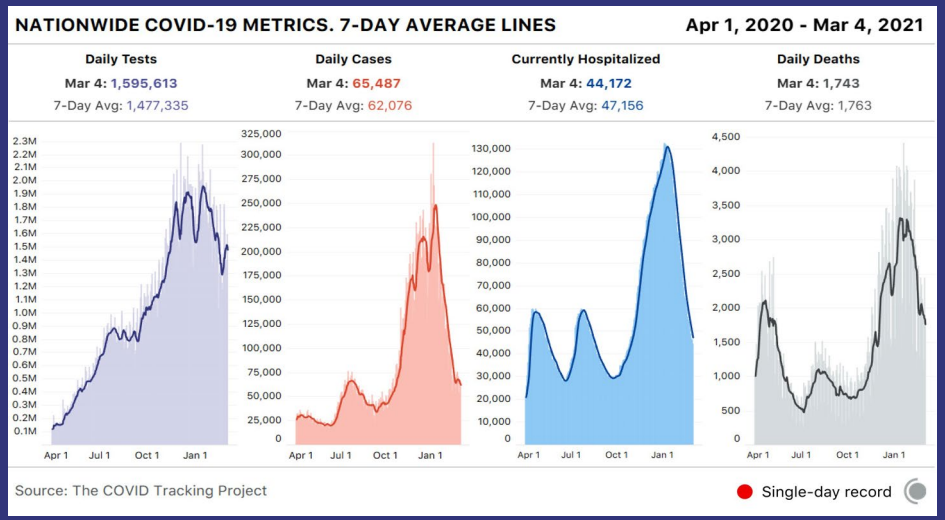


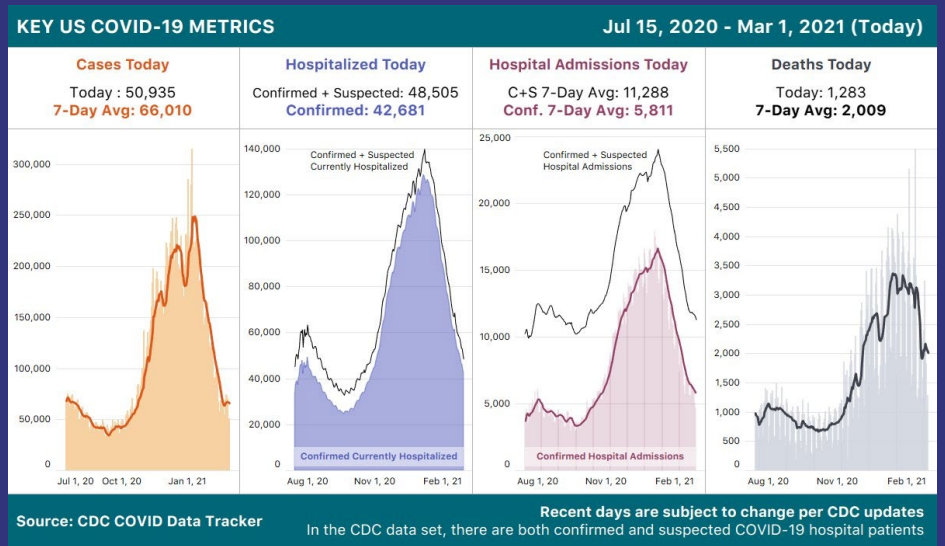
Always include a 7-day or 14-day rolling average to make up for data anomalies

Provide notes and/or annotations with caveats explaining data anomalies

Highlight data points that are data anomalies

Graphic by The New York Times.





The COVID Tracking Project

How to avoid bad takes:

- Understand Dating Schemes
- Study Data Definitions
- Look for Confounding Factors
- Use established relationships between metrics to help guide your interpretation
- Be conservative about what can be known
- Be faithful, not tactical

Increasing Science & Data Literacy Through SciComm

Training data users and the public how to understand COVID-19 metrics:

1. Today's cases represent people who were exposed a week or two ago

Symptoms can take several days to appear. Likewise, it takes time for a patient to get tested, get their results back, and for those results to appear in the dataset. Keep in mind that what you see in daily COVID-19 data is oftentimes a look into the recent past.

2. Weekly trends provide a better picture than any single day's count

Using the 7-day averages can help account for lags in COVID-19 data. Trends are generally more reliable, when you're looking at both local and national data.

3. Watch out for data lags

Each state has their own reporting cadence: some report daily, and some take weekends off. As a result, numbers can have a day-of-the-week effect, which happens when backlogged data shows up a few days later.

Increasing Science & Data Literacy Through SciComm

Training data users and the public how to understand COVID-19 metrics:

4. Hospitalizations are generally more stable than case numbers

Although many hospitalization metrics are inconsistently reported at the state level, national COVID-19 hospitalizations have been one of the best measures of what is actually happening on the ground throughout the pandemic's ebbs and surges this year. It's as close to "real-time" data as we can get.

5. Test positivity is easily skewed

On the surface, the calculation for test positivity (also known as percent positive) is simple: Divide the number of positive tests by the total number of tests, in a select period of time, then multiply the result by 100.Unfortunately, factors like viral prevalence (how many infections there are), testing utilization (who is getting tested), and varying units (the number of people tested or raw tests) can make this metric very difficult to interpret—and even harder to use for public health policy.

The COVID Tracking Project

- Our data was used throughout the US and international media, by researchers, and in public communications by the Trump and Biden administrations.
- We helped persuade dozens of states to provide more complete, comparable COVID-19 data and metadata/data definitions.
- We collaborated with an HHS working group to push for, QA, and signal-boost their release of dramatically better and richer hospitalization and capacity data than has ever been available in the US.
- We responded to thousands of media and general inquiries about government data.