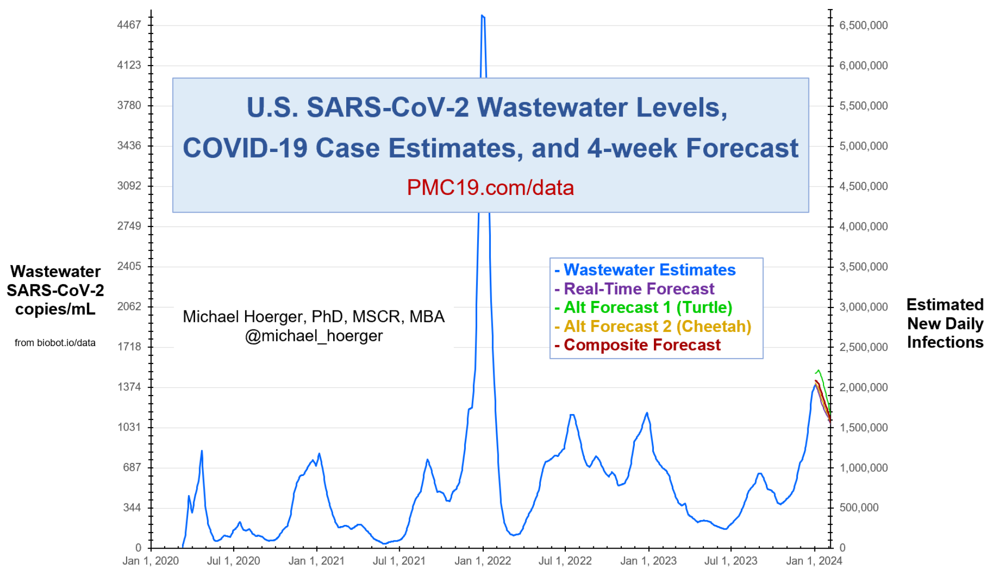
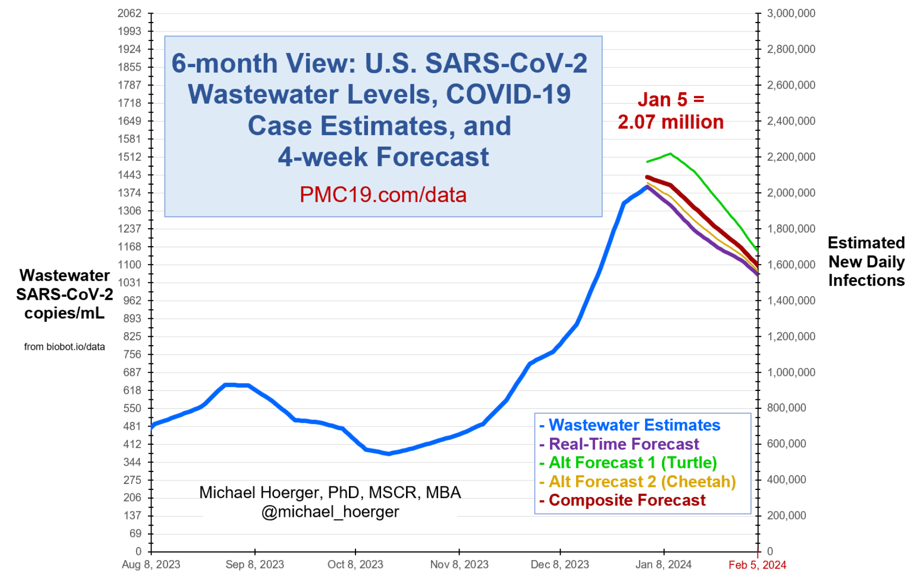
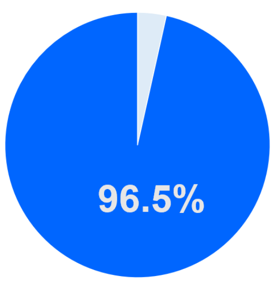
U.S. SARS-CoV-2 Wastewater Levels, COVID-19 Case Estimates, and 4-Week Forecast: Report for January 8, 2024, pmc19.com/data

Michael Hoerger, PhD, MSCR, MBA, Pandemic Mitigation Collaborative







There is more COVID-19 transmission today than during 96.5% of the pandemic.

CURRENT ESTIMATES FOR January 8, 2024
Wastewater Levels (copies/mL)
1413
New Daily Cases
2,056,000
% of Population Infectious
4.3% (1 in 23 people)
New Daily Long COVID Cases
103,000 to 411,000

WEEKLY ESTIMATES FOR January 8, 2024
New Weekly Cases
14,400,000
New Weekly Long COVID Cases
720,000 to 2,878,000

2024 CUMULATIVE ESTIMATES AS OF January 8, 2024
Total 2024 Cases To Date
16,410,770
Total 2024 Long COVID Cases To Date
821,000 to 3,282,000

4-WEEK FORECAST FOR February 5, 2024
Wastewater Levels (copies/mL)
1,096 (-22% lower)
New Daily Cases
1,595,000
% of Population Infectious
3.34% (1 in 30 people)
New Daily Long COVID Cases
80,000 to 319,000

How Does Risk Increase with More Social Contacts?

Number of People	Chances Anyone is Infectious	Number of People	Chances Anyone is Infectious
1	4.3%	25	66.7%
2	8.4%	30	73.3%
3	12.4%	35	78.5%
4	16.1%	40	82.8%
5	19.7%	50	88.9%
6	23.2%	75	96.3%
7	26.5%	100	98.8%
8	29.7%	150	99.9%
9	32.7%	200	>99.9%
10	35.6%	300	>99.9%
15	48.3%	400	>99.9%
20	58.5%	500	>99.9%

Cite as: Hoerger, M. (2024, January 8). *U.S. SARS-CoV-2 wastewater levels, COVID-19 case estimates, and 4-week forecast: Report for January 8, 2024*. Pandemic Mitigation Collaborative. http://www.pmc19.com/data

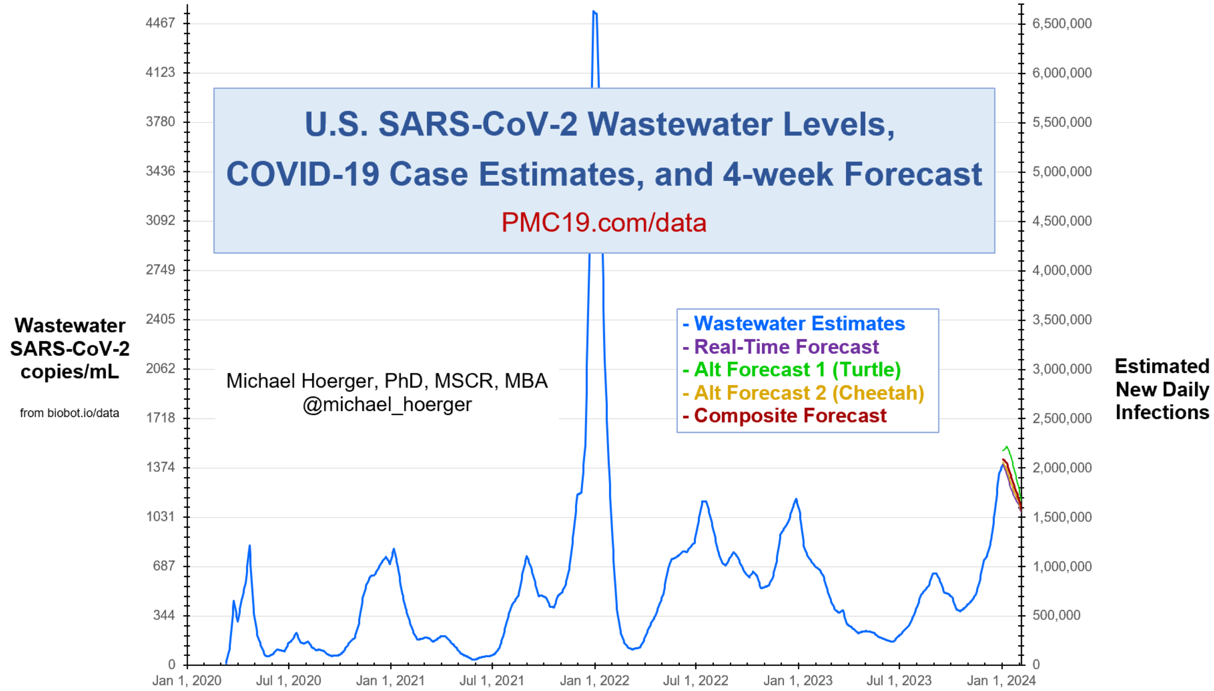
Overview of the Current State of the U.S. COVID-19 Pandemic

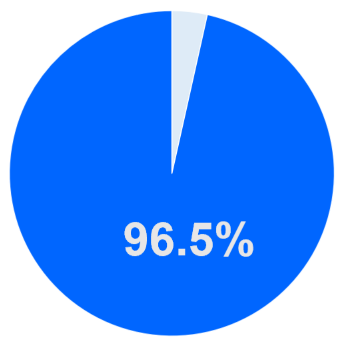
We are in the 8th U.S. COVID wave and 2nd biggest all-time. Any claims that the “pandemic is over” or “COVID is over” run contrary to stated evidence, continue to inflict much harm, and should be labeled misinformation.

U.S. wastewater levels indicate that COVID transmission is higher than during 96.5% of the days of the pandemic and lower than during 3.5% of the days of the pandemic.

- 4.3% (1 in 23) are infectious
- >2 million COVID cases/day
- >100,000 Long Covid cases/day

Weekly new cases stand at 14 million. Weekly Long COVID cases resulting from these infections stand at >700,000.





There is more COVID-19 transmission today than during 96.5% of the pandemic.

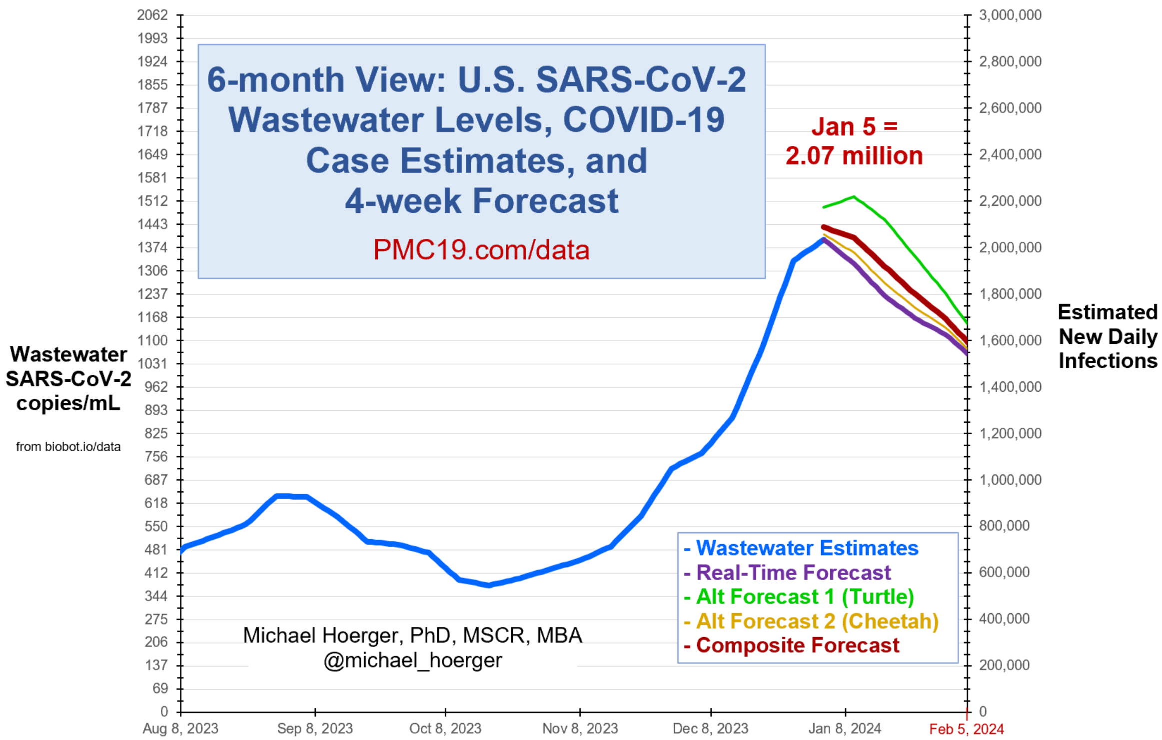
CURRENT ESTIMATES FOR January 8, 2024
Wastewater Levels (copies/mL)
1413
New Daily Cases
2,056,000
% of Population Infectious
4.3% (1 in 23 people)
New Daily Long COVID Cases
103,000 to 411,000

WEEKLY ESTIMATES FOR January 8, 2024
New Weekly Cases
14,400,000
New Weekly Long COVID Cases
720,000 to 2,878,000

2024 CUMULATIVE ESTIMATES AS OF January 8, 2024
Total 2024 Cases To Date
16,410,770
Total 2024 Long COVID Cases To Date
821,000 to 3,282,000

Forecast for the Next Month

We are in the 2nd largest U.S. surge of the pandemic. The different models agree that the peak is just over 2 million U.S. infections/day with >4.2% of the population (1 in 23 people) actively infectious. The composite model suggests the peak was around January 5 (last Friday). If the real-time wastewater levels are accurate (purple line), the peak was on Wednesday, January 3. The past few weeks, wastewater levels have been retroactively corrected upward. The cheetah model (orange) corrects the levels upward by that value, whereas the turtle model (green line) ignores the most recently-reported wastewater level (as downwardly biased) and has a slightly higher peak, not coming until Wednesday, January 10. The red composite line is the average of each model and used for the statistics cited in this report.



CURRENT ESTIMATES FOR January 8, 2024
Wastewater Levels (copies/mL)
1413
New Daily Cases
2,056,000
% of Population Infectious
4.3% (1 in 23 people)
New Daily Long COVID Cases
103,000 to 411,000

4-WEEK FORECAST FOR February 5, 2024
Wastewater Levels (copies/mL)
1,096 (-22% lower)
New Daily Cases
1,595,000
% of Population Infectious
3.34% (1 in 30 people)
New Daily Long COVID Cases
80,000 to 319,000

U.S. Winter 2023-24 COVID Surge - The Peak

	Best Estimate	Range	
Rank among COVID waves	2nd	2nd	2nd
Date of peak	Jan 5	Jan 3	Jan 10
Daily infections at peak	2.07 million/day	2.03 million/day	2.22 million/day
Percentage of population infectious at peak	4.34% (1 in 23)	4.26% (1 in 23)	4.64% (1 in 22)

Magnitude of the Winter 2023-24 Surge

Millions of people in the U.S. will get infected during the winter surge. About 1 in 3 people will get infected during the peak two months and 43% by late February. Those 142 million infections would conservatively translate into an eventual 7 million clinically significant Long Covid cases. Transmission will drop considerably from early February through late March. On February 13, 2.9% of the population will be actively infectious, falling to 2.1% by February 24, 1% by mid-March, and bottoming out around 0.7% in late March. These very long-range projections are more historical medians rather than precise forecasts.

U.S. Winter 2023-24 COVID Surge - The Mountain			
Window	**Approximate Dates**	**Total Infections**	**% of Population Actively Infectious on End Date**
Peak Month	Dec 31 to Jan 29	58 million (about 17% of the population)	3.6% (1 in 28 people) on Jan 29
Peak 2 Months	Dec 16 to Feb 13	107 million (about 32% of the population)	2.9% (1 in 34 people) on Feb 13
Peak 3 Months	Nov 27 to Feb 24	142 million (43% of the population)	2.1% (1 in 47 people) on Feb 24

The CDC has recently approved an updated booster, available to anyone in the U.S. older than 6 months. It is widely available for adults, and now increasingly available for children. Much of the winter surge lies ahead. Yet, vaccine utilization rates this winter remain abysmally low, ranging from 7-36% by state (https://www.cdc.gov/mmwr/volumes/72/wr/mm7251a4.htm). Use the PMC data and local vaccination location information to help people get vaccinated.

Current Risk Based on Number of Social Contacts

This figure shows the chance anyone would be infectious in a group based on group size. For example, in a group of 15-20, there's a 50% chance at least one person is infections. In a group of 50, there's an 89% chance at least one person is infectious. More details for schools are provided on the next page.

How Does Risk Increase with More Social Contacts?

Number of People	Chances Anyone is Infectious	Number of People	Chances Anyone is Infectious
1	4.3%	25	66.7%
2	8.4%	30	73.3%
3	12.4%	35	78.5%
4	16.1%	40	82.8%
5	19.7%	50	88.9%
6	23.2%	75	96.3%
7	26.5%	100	98.8%
8	29.7%	150	99.9%
9	32.7%	200	>99.9%
10	35.6%	300	>99.9%
15	48.3%	400	>99.9%
20	58.5%	500	>99.9%

School-Related Transmission

Risk Table

Schools are often among the highest risk settings for COVID transmission because of the high density of people in classrooms, presenteeism, discontinuation of testing programs, lack of masking requirements, use of ill-fitting and low-quality masks, low vaccination rates, low air cleaning rates, and an emphasis on droplet dogma (handwashing, Lysol) for an airborne virus. As indicated in the letter linked in the next section, the solution is to re-implement comprehensive COVID mitigation, as the U.S. is in the 2nd largest COVID surge of the pandemic.

The following table shows the risk that anyone would have COVID in classrooms, based on varying size, date, and region. The left table shows the chances someone is infectious with Covid based on class size and date. Class size matters considerably. Date does not (for January), as we'll be riding out the peak of the surge, which is more like a month-long plateau. In a class of 8, there's about a 25-30% chance someone would have Covid at any given time in January, absent any precautions using screening, testing, isolation, or quarantine. In a college class of 50, it's more like an 85-90% chance.

Estimated Chances Someone in a U.S. Classroom is Infectious by Week & by Region (pmc19.com model)

Estimates by Date, U.S. Average

Class Size	Jan 1	Jan 8	Jan 15	Jan 22
8	29.1%	29.6%	28.4%	26.9%
16	49.7%	50.5%	48.7%	46.5%
24	64.3%	65.2%	63.3%	60.9%
32	74.7%	75.5%	73.7%	71.4%
40	82.0%	82.8%	81.2%	79.1%
50	88.3%	88.9%	87.6%	85.9%
75	96.0%	96.3%	95.6%	94.7%
100	98.6%	98.8%	98.5%	98.0%
200	99.9%	99.9%	99.9%	99.9%
300	99.9%	99.9%	99.9%	99.9%
500	99.9%	99.9%	99.9%	99.9%

Estimates by Region, Jan 8 Only

Class Size	Northeast & Midwest (Jan 8, Steady)	South & West (Jan 8, Still Rising)
8	39.5%	16.4%
16	63.4%	30.1%
24	77.8%	41.6%
32	86.6%	51.1%
40	91.9%	59.2%
50	95.7%	67.3%
75	99.1%	81.3%
100	99.8%	89.3%
200	99.9%	98.9%
300	99.9%	99.9%
500	99.9%	99.9%

Estimated on Jan 9, using the PMC forecasting model with Biobot data most recently updated in their Jan 9, 2024 dashboard
Pandemic Mitigation Collaborative (PMC), Michael Hoerger, PhD, MSCR, MBA @michael_hoerger

We do not model separately by region. However, if you track Biobot, you'll note levels are higher in the Northeast and Midwest than in the South and West. If we assume the same ratio for Jan 8, this table shows you how the risk varies geographically. The true risk levels could be a bit higher in the South/West if they catch up. The geographic comparisons arguably are not particularly useful because the regional differences tend to be more about the number of specific counties in an extreme surge, while most counties in a region follow the national average. If good local data from Verily or elsewhere, consider that, but otherwise focus more on national numbers. Even using these estimates for the South/West, they are very bad. A class of 32 would have a >50% chance of COVID, absent any screening/testing/isolation/quarantine precautions.

Letter

Dr. Hoerger has prepared a letter for parents seeking to advocate for better COVID mitigation at schools. Parents can consider attaching the letter as an appendix to their own letter. Alternatively, they can replace Dr. Hoerger's letter with their own and use the tables provided, or any graphics from this website. No permission is required to use any of the graphics or data.
https://pmc19.com/data/pmc_school_letter_Dec31_2023.pdf

ASHRAE Air Cleaning Table

Many indoor settings need to increase air cleaning rates to substantially reduce COVID transmission, per the new 2023 national standards for the control of infectious aerosols from the leading organization on indoor air quality, the American Society of Heating, Refrigerating and Air-Conditioning Engineers (ASHRAE). Transmission will be high the next month because a lot of the places that are inherently most risky based on the density of occupants (schools, transportation, restaurants) often also have poor indoor air cleaning. Double whammy. As an example, take schools. Most schools clean the air at 0.8 to 3.0 air changes per hour, substantially lower than ASHRAE 2023 Standard 241. Most school settings should have 6.7 to 9.3 air changes per hour, and sometimes much higher in large densely packed settings.

Simplified Recommendations for Classroom Air Cleaning to Reduce the Risk of Far-Field Airborne Infectious Disease Transmission, Derived from ASHRAE Standards 241 & 62.1

- Mike Hoerger, PhD, MSCR, MBA (Aug 1, 2023)

Educational Facilities	cfm/person (ft^3 / minute per person)	ACH (air changes per hour) for Full-Capacity Rooms
Libraries	40	2.7
Art classroom	40	5.3
Wood and metal shop	40	5.3
Computer lab	40	6.7
Media center	40	6.7
Science labs	40	6.7
University and college labs	40	6.7
Daycare, ages 4 and under	40	6.7
Classroom, ages 5-8	40	6.7
Classroom, ages 9+	40	9.3
Music, theater, and dance	40	9.3
Lecture classroom	40	17.3
Multiuse assembly	40	26.7
Lecture hall, fixed seats	50	50.0

Note. cfm/person based on ASHRAE Standard 241. ACH derived from typical full-capacity person density estimates from ASHRAE Standard 62.1. Use the most comparable educational facility. If half or one-quarter capacity, prorate the ACH estimate accordingly. Calculations focus on reducing risk primarily of far-field transmission in rooms with well-mixed air, and masks remain needed to avoid near-field transmission.

2023 Year in Review

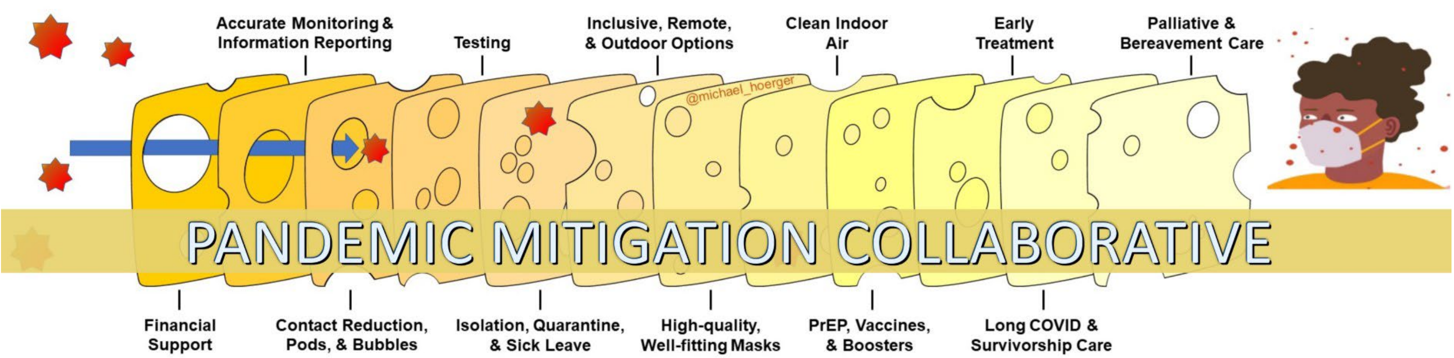
This table documents the toll of the COVID-19 pandemic in the U.S. in 2023. There were over 250 million infections in the U.S. in 2023. If assuming within-year reinfections were 0-30% of total reinfections, an estimated 53-76% of the U.S. population was infected at least once in 2023. With a conservative estimate that 5% of infections will result in clinically significant Long Covid, that's >12 million new Long Covid cases. These outcomes are the product of a highly-risky *laissez-faire* public health response to widespread viral transmission that contradicts the precautionary principle.

CUMULATIVE TOTAL ESTIMATE OF 2023 U.S. SARS-CoV-2 Infections
Total Infections
254,670,431 (53-76% infected)
Total Resulting Long Covid Cases
12,734,000 to 50,934,000

Pandemic Mitigation Collaborative 2023 Advocacy Awards

Congratulations to the PMC Advocacy Awardees!

A PMC committee selected 5 social media accounts we were aware of who were identified as doing an excellent job using the PMC data to advocate for better pandemic precautions on social media, at schools, at workplaces, at home, or with family and friends. We are thankful for everyone using this website to support public health. They were doing plenty of advocacy using PMC data and numerous other sources, so we have merely highlighted one PMC example of their excellent recent work. Ms. Miers has an extensive history of advocacy surrounding Long Covid. Liz O. and Ms. Virgil disseminated valuable information on TikTok. Em & El Organics showed people how to turn an apartment door into a public health dashboard and PPE supply center. Dr. Tran spruced up PMC graphics and shared widely across social media platforms, especially Instagram. We singled out Ms. Virgil's work for a "Top Commendation" as she devoted substantial time burning the midnight oil get an engaging and informative TikTok video out right the day it become clear we were headed into the 2nd biggest surge of the pandemic.



PMC 2023 Advocacy Awards

Laura Miers (Twitter: @LauraMiers)

Liz O (Twitter: @BettyB2007)

Em & El Organics (Twitter: @EmAndElOrganics)

Dr. Lucky Tran (Twitter: @luckytran)

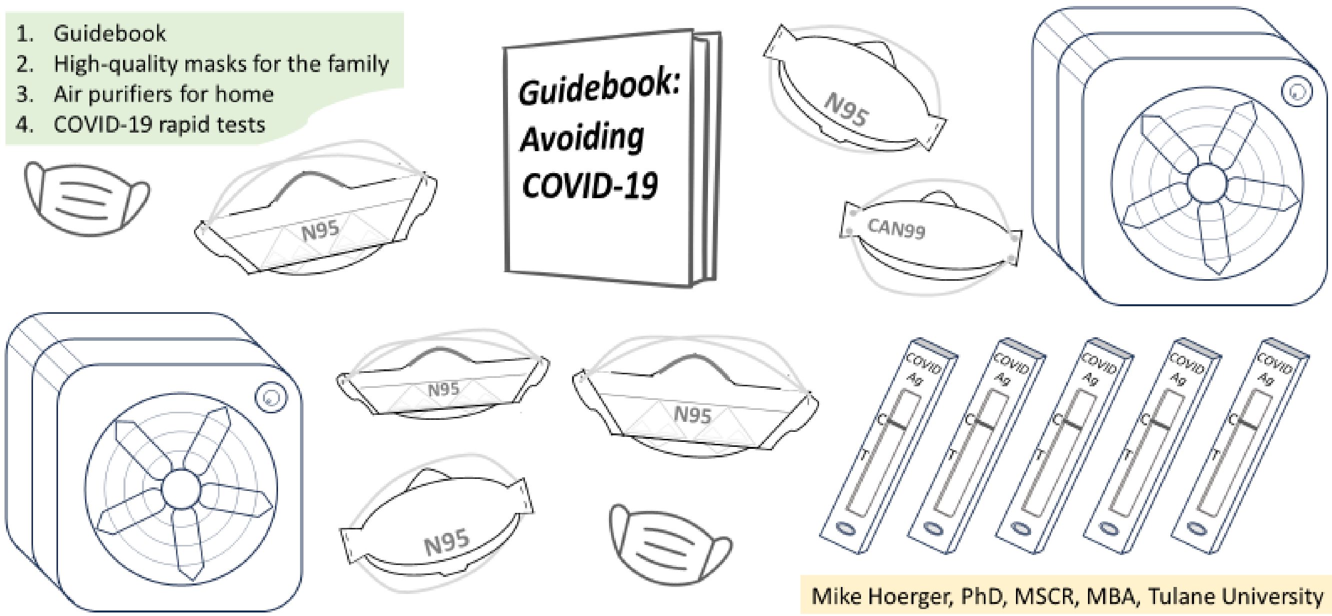
Christina Virgil (Twitter: @ChristinaVirgil)

Examples of their work is summarized here: https://twitter.com/michael_hoerger/status/1744580904356401284

Announcement - 2024 Pandemic Research Grant Submission: Calling All Scientists, Clinicians, Patients, and Other Stakeholders!

We are preparing a grant application for a research project aimed at helping U.S. adults with cancer to avoid COVID infections. A letter of intent is due shortly, and the grant will get submitted in May. We are seeking a team of scientists, clinicians, patients, and other stakeholders passionate about COVID mitigation who may want to learn more about and potentially get involved with this project. Please help!

The project builds on our ongoing pilot projects funded by Tulane University and the American Cancer Society.



Please complete this online interest form, and we will follow up with details.
https://tulane.co1.qualtrics.com/jfe/form/SV_3CwpzKDXSefD0uq

We plan to host an informational meeting in the near future. Follow up with Dr. Hoerger by email too as desired.

Trustworthy Health Evidence

Transparency

The pandemic has reminded us that trust is essential in public health, and we have made every effort to provide transparency. Every PMC report remains posted publicly. A methodology video describes the gist of the analytic approaches used, with enough detail that others can construct similar models. Dr. Hoerger's biography and publication record are publicly available. Many of those making similar dashboards and forecasts work in private industry or wish to avoid scrutiny so do not post information publicly or do so using anonymous accounts or pseudonyms.

Lack of Conflicts of Interest

These reports are written by Dr. Hoerger who has no financial conflicts of interest. This means no ownership of equity stakes or individual stocks. Dr. Hoerger actively avoids funding from the pharmaceutical industry and other private corporations. He has no investment in mask, HEPA, or other mitigation companies. As a scientist, he is required to complete financial forms as a part of being a university professor as well as when publishing at top medical journals, like JAMA Oncology. An underappreciated aspect of financial conflicts of interest is that clinicians and scientists sometimes gain substantial revenue from social media advertisements (e.g., YouTube, Twitter) and newsletter subscriptions (e.g., Substack, Patreon); this can lead to significance "audience capture" in which the scientist gains increasing financial remuneration for telling an audience what they want to hear, and this is a huge problem among those posting disinformation online. Dr. Hoerger intentionally gains no revenue through this website, social media, nor other platforms, as doing so would undermine his scientific credibility. The closest PMC collaborators are the co-authors and co-investigators noted in the various projects and publications on the PMC homepage, and there are no noteworthy conflicts of interest there either.

Documented Evidence of Accuracy

The PMC Forecasting model can be conceptualized as "participatory action research," meaning that it is public health activism first, iterates and improves with feedback, and is research second. An alternative approach would be to publish a model first and then implement, but that would be a horribly unsuccessful strategy. Using historical data, our forecasting model explains 96% of the variance in 1-week forecasts, which is exceptional. However, implementing forecasts in real-time adds new challenges. Wastewater levels are reported with inconsistent cadence and often corrected retroactively, so forecasting in real-time is more challenging, meaning that real-world estimates of accuracy are more important than those derived from a historical database. Moreover, it has been a learning experience understanding what metrics matter most to the public, and these metrics vary across different points of the pandemic. Accordingly, all reports are public so that users can evaluate accuracy themselves in real-time. However, a scientific report will ultimately document accuracy. Accuracy levels will guide a 2.0 model of the forecast likely to launch in spring or summer of 2024, which may include the same or different alternative forecasts, and add confidence intervals based on historic levels of accuracy.

A useful way to evaluate accuracy is to take one of our public reports and then compare it with what unfolds four weeks later. We believe you will be impressed with the results and that any discrepancies are minor and not of public health significance, given that much of the public is magnitudes off in their subjective estimates of transmission. As an example, in our December 25 report, we noted that our first-posted Christmas Day risk table (posted October 30, 2023, with that report cited in JAMA Oncology) was quite accurate in characterizing the eventual risk on Christmas Day (noted in the December 25, 2023 report, though Biobot may retroactively adjust wastewater levels marginally).

Posted on December 25

In the U.S., What's the COVID Risk for Christmas Day?

Number of People	Chances Anyone is Infectious	Number of People	Chances Anyone is Infectious
1	3.5%	25	58.6%
2	6.8%	30	65.3%
3	10.0%	35	70.9%
4	13.2%	40	75.6%
5	16.2%	50	82.8%
6	19.1%	75	92.9%
7	21.9%	100	97.1%
8	24.6%	150	99.5%
9	27.2%	200	99.9%
10	29.7%	300	>99.9%
15	41.1%	400	>99.9%
20	50.6%	500	>99.9%

@michael_hoerger Estimated December 25, 2023

Posted on October 30

In the U.S., What's the COVID Risk for Christmas Day?

Number of People	Chances Anyone is Infectious	Number of People	Chances Anyone is Infectious
1	3.1%	25	54.8%
2	6.2%	30	61.5%
3	9.1%	35	67.1%
4	11.9%	40	72.0%
5	14.7%	50	79.6%
6	17.4%	75	90.8%
7	19.9%	100	95.8%
8	22.5%	150	99.1%
9	24.9%	200	99.8%
10	27.2%	300	>99.9%
15	37.9%	400	>99.9%
20	47.0%	500	>99.9%

@michael_hoerger Estimated October 30, 2023

Negligible differences between the long-range forecast and today's table. In a group of 20, for example, the forecasted risk was 47%, and today we estimate it at 50.6%. Hopefully this helped with planning.

Documented Evidence of Validity Relative to Other Data Sources

The public has genuine reasons to ask questions about wastewater-derived case estimates and forecasts. These concerns fall into two main themes. One, much of the general public has observed how important metrics (reported case counts, positivity ratio) have gone from valuable to fallible as testing has declined and largely switched from reported PCR tests to unreported home rapid tests. Two, with each new subvariant, some also wonder whether wastewater data could over or underestimate transmission. With regard to the fallibility of testing-based data, a tremendous benefit of wastewater is that it does not rely on self-report. With regard to subvariants, it is important to know that the process of converting raw wastewater into useful, normalized metrics of viral levels is oversee by wastewater scientists with specialized expertise in addressing our most obvious concerns (e.g., new subvariants, regional differences in demographics, changes in rain water). It is a good example of the Dunning-Kruger Effect for anyone to assume they have more expertise than the scientific experts who translate sewage into viral level estimates, but our public health infrastructure fosters much skepticism. Moreover, by outsourcing wastewater measurement to third parties who must compete with one another, methodology surrounding wastewater quantification remains proprietary to protect competitive interests, and this reduces transparency and trust. Fortunately, there are two easy checks on wastewater validity.

Validity Check #1: Independent Cross-Cultural Comparisons. If the PMC model provides a useful estimate of cases, it should converge with other independent estimates. Unfortunately, other U.S. based models are private or anonymous, so not appropriate comparisons. Nonetheless, there are high-quality estimates outside the U.S. For example, Dr. Moriarty's lab uses wastewater data in Canada to make similar estimates, and the U.K. runs a strong testing surveillance program to assess case rates. Despite being from independent groups, using slightly to considerably different methodology, and in different countries, all of the estimates are within the same ballpark. The PMC estimates are reasonable given what is known of similar estimates from other countries with high-quality data. Note that both of the non-U.S. estimates are a bit outdated given holiday-related reporting lags.

International COVID Statistics from High-Quality Data Sources

Compiled by the Pandemic Mitigation Collaborative (PMC19.com), January 8, 2024

	U.S.A.	Canada	U.K.
% of Population Actively Infectious	**4.3% (1 in 23)**	**5.0% (1 in 20)**	**4.2% (1 in 24)**
Chances Someone is Infectious…			
In a Group of 10	**35.6%**	**40.1%**	**34.9%**
In a Group of 30	**73.3%**	**78.5%**	**72.4%**
In a Group of 50	**88.9%**	**92.3%**	**88.3%**
Primary Data Source	Wastewater	Wastewater	Surveillance Testing
Reference	Michael Hoerger, PhD, MSCR, MBA @michael_hoerger pmc19.com/data	Tara Moriarty, PhD @MoriartyLab covid19resources.ca	Alex Glaser, PhD UK Health Service Agency tinyurl.com/pmcukhsa
Last Updated	January 8, 2024	December 17, 2023	December 21, 2023

Validity Check #2: Examining Consistency Relative to Other Metrics. For the past 18+ months, the relationship between self-reported cases and estimates of true cases has remained remarkably consistent across various indicators of each variable. A simple heuristic is that there are 25 actual cases for every 1 self-reported case tracked in federal statistics. This fluctuates a bit depending on whether in a lull, rising surge, or falling surge, often in the 15-30x range. Simply review IHME self-reported versus actual case data, available through April 1, 2023 (https://covid19.healthdata.org/united-states-of-america?view=infections-testing&tab=trend&test=infections). The same is true for our wastewater-derived estimates of true cases, relative to self-reports of actual cases provided by states and aggregated nationally by BNO News (https://twitter.com/BNONews) and others. If this relationship were to decouple suddenly, such that estimates of true cases were 100x or 1,000x that of self-reported cases, that would suggest that a particular subvariant may be disrupting wastewater estimates in ways that were outsmarting wastewater scientists, but there is no evidence for such a claim at present.

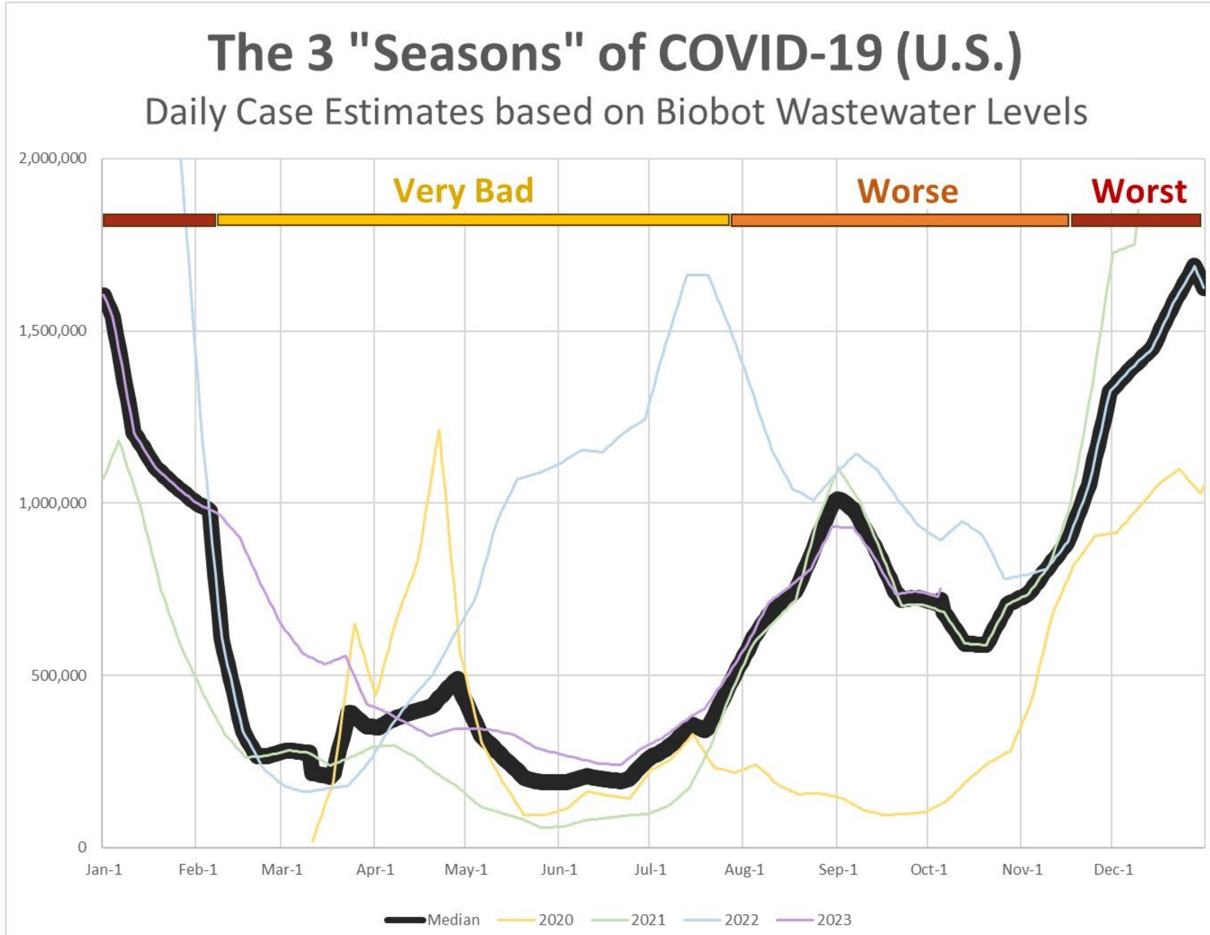
PMC Trusted in Peer-Reviewed Scientific Journals, News Media, and Funded Grant Applications

Examples:

- JAMA Oncology: https://jamanetwork.com/journals/jamaoncology/fullarticle/2813585
- BMC Public Health: https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-023-16787-1
- TODAY: https://www.today.com/health/news/covid-wave-2024-rcna132529
- Forbes: https://www.forbes.com/sites/judystone/2023/12/01/cdc-improves-their-covid-19-reporting-with-a-new-wastewater-dashboard
- Salon: https://www.salon.com/2023/10/19/a-lapse-in-wastewater-detection-is-worrying-scientists-about-distorted-data/
- Grants funded intramurally by Tulane University and externally by the American Cancer Society

Forecast for the Longer-Term – Annual Trends

*This section is being reshared from a prior report, as some people may need to make decisions about travel, surgery, and other important events several months out.



Is COVID-19 "seasonal"?

Not in any meaningful sense of the word. The following graph uses historical Biobot #wastewater data to estimate daily case rates using the PMC model. Thin lines show 2020 (yellow), 2021 (green), 2022 (blue) and 2023 (purple). The black line is the median. It is not really a forecast, but merely a summary of historical data. To the extent the median provides a reasonable approximation of the future, it is a useful starting point for a gift-level forecast.

Season 1: "Very Bad Transmission"

Focusing on the median, you'll see that case rates tend to be lower (but still in the 250-500K/day range) from mid-Feb through the end of July. These are valuable data. If I needed to schedule a non-urgent surgery, when would I do it? Late February, when transmission has often dropped, but before the general public not monitoring wastewater has realized so, perhaps meaning some people are still using airborne precautions. You're basically hopefully beating the transmission "market." I'd also be prepared to cancel an appointment or push back 6 weeks if needed.

Season 2: "Worse Transmission"

Again, focusing on the median, you see a late summer wave from August through mid-October. This is the clearest indication that C19 is not "seasonal," if people are using that term to mean an annual event. If we were doing two boosters a year, it seems like booster 1 would roll out in July. Why do we have this wave? Schools have very little mitigation (poor air quality, little/no testing, little/no masking, low vax rates). Also, the fleeting immunity from winter boosters and infections has waned. If I had an urgent maskless medical/dental visit, I'd schedule mid-October through early November and cross my fingers (around my HEPA). It's still high transmission but about to get worse. This is also a good time to stock up on N95s, rapid tests, and HEPA filters before the prices may increase, scarcity may become a problem, or one has an infection in the home. Travel insurance is wise.

Season 3: "Worst Transmission"

From mid-November to mid-Feb, transmission is extremely problematic, according to the median line. Everybody should be wearing high-quality masks, testing as frequently as possible, improving indoor air quality, and moving activities outdoors and remote.

A Couple Caveats

Seasonality. Some people use the word "seasonal" to mean predictable, rather than merely a discrete 2-3 month season of transmission. In some ways, transmission is predictable. You'll see the 2023 purple line has followed the median very closely. However, we're talking about a very small sample size of years, so one would expect one of the years to mostly follow the median. Also, there are clear discrepant cases. BA.1 goes off the chart (winter 2021 to early 2022). The 2022 summer wave was also sizable. My approach is to make longer-term plans based on the median line and then be prepared to shift plans toward more remote activities if a large wave picks up. Hopefully, transmission becomes more predictable as years go by, but I'm not betting on it yet.

Case estimates. If you have followed the PMC dashboard, you'll know these are estimated by linking Biobot wastewater levels to IHME true case estimates. I would find case estimates 15% higher or 30% lower also reasonable and discuss these estimates with many modeling experts. There are also some more sophisticated models, where I believe an argument can be made that waves are actually marginally more leptokurtic (spikier mountains and deeper valleys than shown here).

General Technical Notes, Not Specific to the Current Week's Report

Status of Biobot wastewater reporting

The estimates and forecast described here use wastewater data reported by Biobot. Biobot is now updating their data on Fridays or Mondays, and the CDC has awarded several prior Biobot sites to a company called Verily. The transitionary phase at Biobot seems mostly through, though Biobot is contesting the contract reassignment in court. As long as national wastewater data are being reported, the PMC reports will continue.

Case estimates

Case estimates were used by evaluating various potential multipliers to go from wastewater levels to cases. To identify true cases, not merely just reported cases, I used the IHME's case estimates for January 1, 2021 through April 1, 2023 (https://covid19.healthdata.org/united-states-of-america?view=cumulative-deaths&tab=trend). I compared wastewater with their case estimates on the 1st of each month. The correlation was r=.94. The maximum possible correlation is 1.00, so that is freakishly high, higher than just about any of the 10,000 or so correlations I've ever run. I was hoping for a correlation of r=.70 or higher, which still would have been great. Basically, wastewater is a supreme indicator of case rates. Next, I examined multipliers. Are cases 10x the arbitrary wastewater metric? 10,000x? Something else? Take cases and divide by wastewater at each data point, then find a summary metric (mean, median, trimmed mean, etc.). The metric I found most defensible was to use a +/-10% trimmed mean (average that excludes extreme data points, where case estimates are more error-prone), where each unit of wastewater translated into 1455 cases. I would find multipliers of 1000 to 1700 (31% lower to 17% higher) also reasonable. Arguably, case rates are magnitudes (10-100 times) higher than many people expect, so these details have minimal practical significance for everyday decision making. There are also more sophisticated strategies, such as regression models, but I found those results to be counter-intuitive (e.g., positive intercept, where I would have expected zero or negative). One can set the intercept to zero, use various heteroscedasticity-related techniques, and correct for the lack of imperfect reliability, but most of that is over the heads of people using this model and would accomplish little more than the trimmed multiplier method. The multiplier method has also led to techniques (only posted on Twitter thus far) for making regional estimates using very simple multipliers. Elegant is good.

Percentage infectious

After estimating the current number of new infections, it is relatively straightforward to estimate the percentage of the U.S. population actively infectious with COVID-19, but there are several caveats worth noting. One, the U.S. population is assumed to be 334,565,848. This was the CDC-estimated U.S. population on the final day of the IHME case estimation model. The number of new daily cases divided by the population tells one the percentage of the population newly infected today, often small at around 0.3% or less. Two, consider the infectious window. The percentage of the population infectious depends on the percentage of new people infected but also the duration people stay infectious. The model assumes people stay infectious for 7 days. Low estimates are that people are infectious for an average of 5 days (this defies the preponderance of the evidence, in my view), and high estimates are more like 10 days (too high in my view, based on a preference for round numbers). Other compelling estimates are more like 8-8.5 days. This duration may change over time, based on new variants, new vaccines, vaccine utilization rates, and treatments. If assuming the infectiousness duration is 10% longer, multiply by 1.10. If assuming 20% shorter, multiply by 0.80. New cases divided by the population equals new daily infections. Note also, these are merely averages and do not reflect individual variation, as some get infected and are not contagious, whereas others get infected and remain infectious likely for months (extremely rare). New daily infections multiplied by the number of days infectious indicates the percentage of the population actively infectious.

Long COVID

Long COVID case estimation. The lower and upper bounds for Long COVID case estimates assume that 5-20% of people infected with SARS-CoV-2 will develop Long COVID as a result of that infection. Some published reports and analysts have suggested lower (1%) or higher (40%) values. A useful framework for thinking about these estimates is that the low value is more indicative of people experiencing serious, enduring, known harms, whereas the upper estimates are closer to the number experiencing disruptive symptoms for at least several months, perhaps with full or partial recovery. These estimates do not indicate unknown long-term harms. For example, if infections increase the risk of cancer or cardiovascular disease substantially and with increasing risk over 10-30 years, that is not captured well by these metrics. The metrics also do not encompass the 1.2 to 1.8 million Americans who have died of COVID-19. Future models may incorporate estimates of mortality. Finally, the estimates project the number who will ultimately experience Long COVID from a new infection, but that is several months down the line. The estimates reflect future implications. For simplicity of interpretation, they are not modeling the number of new Long COVID cases today that resulted from infections three months ago.

General forecasting model specification

The forecasting models are elegant, meaning simple and effective. In regression analyses using historical pandemic wastewater data, the model explains 96% of the variance in the following week's forecast. The model is simple. It includes the year (2020, 2021, 2022, or 2023). It includes the historical median (switched from average on 12/11/23) for the current half month; imagine the year sliced into 26 pieces, and it incorporates data on the historical median for that half month (e.g., second half of September). The model also incorporates four lagged variables, the wastewater levels 1, 2, 3, and 4 weeks ago. Overall, you can think of the model as having two main processes. One incorporates what we know historically. The other incorporates what has been happening the past several weeks. The historical data are useful because transmission mostly, but not always, follows a particular monthly pattern. It is not seasonal in that there are not just three bad months a year, but there is month-to-month variation, and sometimes even useful differences between the first versus second half of the month. The use of recent wastewater estimates helps in several ways. It lets the model know if something about the current point in time differs dramatically from the historical data, and it quickly adapts the model to changes, such as if a wave is starting or ending,

Real-time model (purple line)

This model assumes that real-time data reports of wastewater levels are accurate. However, real-time data often get corrected. Some sites may be slow reporting, and if there is a bias built in, such as places with high transmission being late to report, that would be a problem. Often, the real-time reports are quite accurate, but occasionally they have been corrected substantially a week later. The general model places a lot of weight on the most recent data, so any errors here can lead the model to assume a wave is picking up that really is not (false alarm) or that things are improving better than expected (false hope).

Alt model #1, turtle (green line)

The turtle model moves slow and steady. It completely ignores the most recent week's worth of data from Biobot, treating it as unreliable. It will ignore false fluctuations inferred from inaccurate real-time reporting. However, it will be slower to respond to real changes, such as the onset in a new wave or the decline in a wave that has peaked.

Alt model #2, cheetah (orange line)

The cheetah model moves fast. It aims to correct for biases in real-time data reports. If last week's real-time report overestimated levels by 10% upon correction, it assumes this week's real-time report suffers the same bias. If last week's real-time report underestimated true levels, it assumes the same for this week. If last week's real-time report was accurate, it will look similar to the real-time model. This model is very good if there is a bias, such as if areas with high transmission experience delays in reporting. However, it can also be overreactive. If there was some error in a real-time report that was just "random" rather than biased in a particular correction, it will tend to overcorrect the next week's model.

Composite Model (red line)

This is the arithmetic average of the three models. It's what's used for deriving all of the statistics reported. When all of the individual models are very close to the average, that suggests high confidence. When the models make vastly different predictions, that suggests more uncertainty in the data, largely based on perceptions of the accuracy of real-time wastewater reporting.