Incorporating Crowdsourced Perceptions of Human Behavior into Computational Forecasts of US National Incident Cases of COVID-19

David Braun[1]*, Daniel Ingram[2], David Ingram[2], Bilal Khan[3], Jessecae Marsh[1], and Thomas McAndrew[3]

[1] *Department of Psychology, Lehigh University, Bethlehem, PA, US*
[2] *Actuarial Risk Management, LTD, Austin, TX, US*
[3] *College of Health, Lehigh University, Bethlehem, PA, US*

Computational UncertaintyLab

* *Corresponding author: dab414@lehigh.edu*

Introduction

Machine vs. Human Predictions

- The majority of forecasts of an infectious disease are generated by **machine predictions** (ie, computational models). However, computational models require large amounts of data to train on, which is often unavailable after a sudden outbreak of an infectious disease.
- In contrast, **human predictions** are rapidly available and past research has shown that crowd sourced human judgments can accurately predict many phenomena[3,5], including infectious agents[1,2,4,6].

Direct vs. Indirect Predictions

- **Direct predictions** are collected by asking humans to estimate the probability of a future event of interest. Direct predictions take advantage of a human's ability to incorporate into predictions information from structured data as well as subjective information, intuition, and expertise.
- **Indirect predictions** can be collected by asking a crowd about covariates that may be related to the target of interest. Indirect predictions offer an opportunity to train a statistical model on both measured, objective data and subjective data.

Approach

- Distribute a survey to a nationally representative sample asking questions about the extent to which their community is adhering to CDC's non-pharmaceutical COVID-19 interventions.
- Leverage survey responses to predict nation-wide incident COVID-19 cases up to four weeks in the future.

Research Question:

Can crowdsourced, indirect predictions be a useful signal for forecasting incident cases of COVID-19?

Method

Survey and Participants

- Distributed between August 30, 2020 and April 28, 2021.
- 21 questions; Distributed on Pollfish and SurveyMonkey.
- Response options ranged from 1 (no adherence) to 5 (full compliance).
- N = 10,852 (both paid and volunteer).

Example Questions
What percent of people in your community do you notice are usually:
Wearing a mask in public?
Maintaining social distance?
Staying at home?

Mean Perceived Adherence (MEPA)

- Mean perceived adherence (MEPA) is defined for a specific question q and at a specific time t as the average of $x_{t,i,q}$ over participants (P), or:

$$\text{MEPA}_{q,t} = N^{-1} \sum_{i=1}^{P} x_{t,i,q}$$

- MEPA is intended to measure an aggregated adherence to a specific type of non-pharmaceutical intervention.

Analysis Approaches

Question	Analysis	Detail
How did people respond to the survey?	Plot proportion of missing responses by question; plot expected vs. observed responses by state.	Proportion of missing responses by question might reveal how easy or difficult the questions were to answer; response rates by state will reveal whether our sample was representative of the US population.
Did perceptions of adherence across different survey questions change similarly over time?	Fit a hierarchical clustering algorithm to all 21 MEPA time series for 2 through 10 clusters.	Dissimilarity between two time series was computed using the Euclidean distance. The Silhouette coefficient was used to assess the quality of fitting all clusters.
Is perceived adherence associated with incident COVID-19 cases?	Correlate MEPA with 0-4 week ahead incident COVID-19 cases.	A Pearson correlation coefficient was calculated between MEPA at epiweek t and US national incident cases at: epiweek t, epiweek $t+1$, $t+2$, $t+3$, and $t+4$.
Does perceived adherence improve probabilistic forecasts of incident COVID-19 cases?	Fit an SIR and VARIMA model to incident cases and assess improvement added by MEPA.	An SIR model was fit to the timeseries of incident cases, and a VARIMA model was fit to the residuals either with or without a MEPA time series. Forecasts were scored using the weighted interval score (WIS).

Results

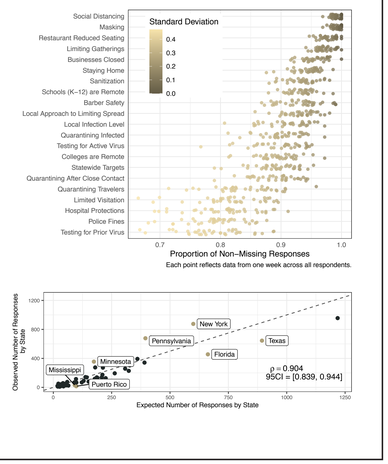
How did people respond to the survey?

People were more likely to respond when asked about easily perceivable compliance actions.

- Higher response rates for actions such as **social distancing** or **masking** than about actions such as whether violations of COVID-19 restrictions result in **fines or police enforcement** or whether people are **getting antibody testing to detect prior virus infection**.
- These response rates suggest people are sensitive to what they can and cannot estimate based on observation.

People's responses were geographically representative of the US population.

- **Expected number of responses by state** was calculated as the proportion of a state's population to the US population.
- States like Pennsylvania and New York were oversampled due to convenience sampling.



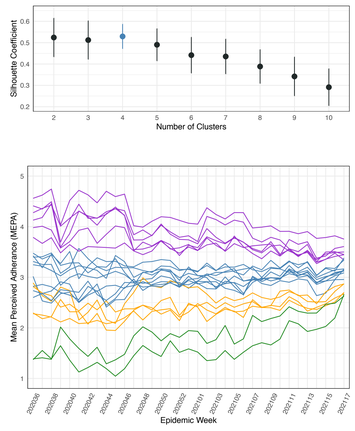
Did perceptions of adherence across different survey questions change similarly over time?

Perceptions of adherence across survey questions grouped into four clusters.

- A hierarchical clustering algorithm conducted on the MEPA time series for all questions using the Silhouette coefficient to assess goodness of fit determined that four clusters best characterized the data.

Perceptions of adherence either steadily increased, decreased, or remained moderate.

- Perceptions of adherence trended toward moderate over the course of the survey period. This could be due to changing perceptions or increased sample sizes over time.
- That survey questions could be clustered suggests a shorter survey might have sufficed. Although how the content among clustered questions is related is unclear.



Is perceived adherence associated with incident COVID-19 cases?

Several types of adherence perceptions were strongly associated with incident COVID-19 cases.

- Several of the more easily perceivable types of adherence actions—such as whether **schools are remote**, **restaurants have reduced seating**, and people are **social distancing**—were strongly associated with incident COVID-19 cases.
- Perhaps surprising that even though it is easily perceivable, perceptions of whether people are **masking** had no strong association with incident COVID-19 cases.

Causal relationships between adherence perceptions and incident COVID-19 cases are unclear.

- It's plausible that increased **social distancing** is causing fewer incident cases, and it's also plausible that **colleges being remote** is being caused by increased incident cases.

* 95% CI does not include 0.

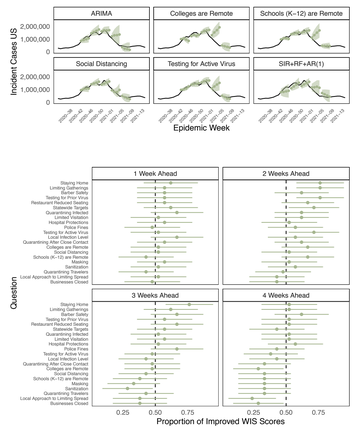
Does perceived adherence improve probabilistic forecasts of incident COVID-19 cases?

Forecasts were most accurate when including perceived adherence across all questions into the model.

- The most accurate forecast of incident COVID-19 cases (panel F) involved first fitting an **SIR model** and then fitting a **random forecast** to residuals that included **MEPA values for all questions** plus an **AR(1) model**.

The majority of perceived adherence question types improve one and two week ahead forecasts of incident cases.

- The figure to the right plots the proportion and 95CI of WIS scores that were improved for an SIR plus VARMA model including MEPA time series to a control model not including MEPA time series.
- Fewer MEPA time series improve three week ahead forecasts, and four week ahead forecasts are improved only modestly.



Discussion

Advantages and disadvantages of human judgment.

- **Advantages** of human judgment include that it is fast to collect and that humans have access to information not available to computational models, such as intuition and subjective observation.
- **Disadvantages** of human judgment include that people are susceptible to biases that can be triggered by subtle changes in how a judgment prompt is presented.

Future research should match the spatial scale of judgments and incident cases.

- In the present study, community-level perceptions were leveraged to predict national-level incident cases.
- Much stronger connections may be observed between state- or community-level judgments and state- or community-level incident cases.
- Reliable predictions at the community level could be a powerful tool for community leaders and policy makers.

Conclusion:

Crowdsourced, indirect predictions *can* be a useful signal for forecasting incident cases of COVID-19.

Acknowledgments & References

Acknowledgments

We thank members of the COVID Mitigation Monitoring Project (CMMP): Dale Hall, FSA, CERA, MAAA, CFA, Robert Wolf, FSA, CFA, CERA, Kailan Shang, FSA, CFA, PRM, SCJP, and John Stark, FSA, CERA. We thank the Society of Actuaries for the support of the data collection. We also wish to thank all participants who offered their perceptions over the survey period.

References

[1] Nikos I Bosse, Sam Abbott, Johannes Bracher, Habakuk Hain, Billy J Quilty, Mark JK Edwin van Leeuwen, Anne Cori, Sebastian Funk, et al. Comparing human and model-based forecasts of covid-19 in germany and poland. *medRxiv*, 2021.

[2] David C Farrow, Logan C Brooks, Sangwon Hyun, Ryan J Tibshirani, Donald S Burke, and Roni Rosenfeld. A human judgment approach to epidemiological forecasting. *PLoS computational biology*, 13(3):e1005248, 2017.

[3] Anca M Hanea, Victoria Hemming, and Gabriela F Nane. Uncertainty quantification with experts: present status and research needs. *Risk Analysis*, 2021.

[4] Thomas McAndrew and Nicholas G Reich. An expert judgment model to predict early stages of the covid-19 outbreak in the united states. *Medrxiv*, 2020.

[5] Thomas McAndrew, Nutcha Wattanachit, Graham C Gibson, and Nicholas G Reich. Aggregating predictions from experts: A review of statistical methods, experiments, and applications. *Wiley Interdisciplinary Reviews: Computational Statistics*, 13(2):e1514, 2021.

[6] Gabriel Recchia, Alexandra LJ Freeman, and David Spiegelhalter. How well did experts and laypeople forecast the size of the covid-19 pandemic? *PloS one*, 16(5):e0250935, 2021.

This poster is available online at: https://davebraun.net/midas