

Quantifying the Effects of COVID-19 on Mental Health Support Forums

Laura Biester*, Katie Matton*, Janarthanam Rajendran,

Emily Mower Provost, Rada Mihalcea

Computer Science & Engineering, University of Michigan, USA

* = equal contribution

Motivation

COVID-19 has negative consequences for mental health



Image: eamesBot/Shutterstock

Objective

We aim to quantify how COVID-19 has impacted the behavior of people who post in mental health support forums on Reddit.

Research questions:

1. How often is COVID-19 discussed in mental health subreddits?
2. Has COVID-19 changed the number of users posting in mental health subreddits?
3. Has COVID-19 led to changes in the discussions that users have surrounding mental health?



Data

Reddit posts from three mental health subreddits from January 2017 – May 2020

	r/Anxiety	r/depression	r/SuicideWatch
Year	Mean Daily Post Count	Mean Daily Post Count	Mean Daily Post Count
2017	95	279	91
2018	164	449	188
2019	211	622	285
2020	243	618	370

Mean Daily Post Count by Year

Anyone go years without knowing what they were experiencing was an anxiety disorder?

I always thought anxiety was all physical symptoms... But I realize now my symptoms include convoluted thoughts about people and situations, genuinely thinking I'm a bother to others... the list goes on.

....

I wanted to see if anyone has had similar experiences of being slow about coming to terms with their anxiety?

Paraphrased
r/Anxiety
Reddit Post
Example

Approach Overview

1. Split data into **pre-COVID** and **post-COVID** periods
2. **Compute daily metrics** from posts in each period
3. Use **time-series intervention analysis** to measure how metrics in post-COVID period compare to what is expected based on pre-COVID data

Pre/post COVID period

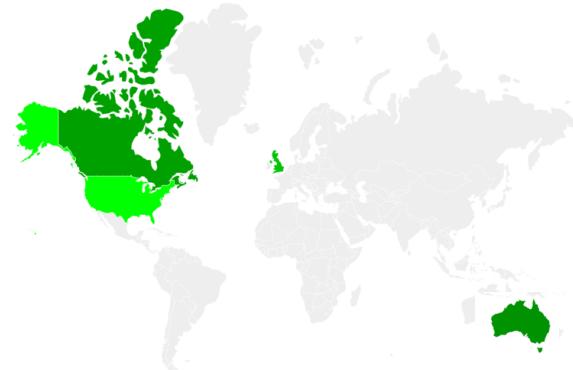
We split the data into two time periods:

Pre-COVID:

Jan 1, 2017 – Feb 29, 2020

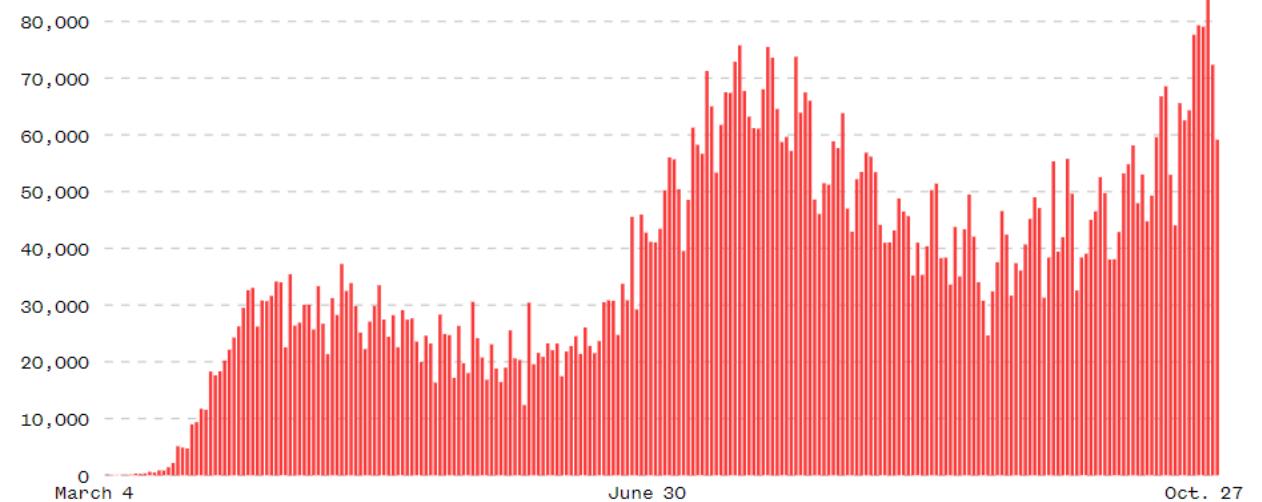
Post-COVID:

March 1, 2020 – May 31, 2020



~50% of Reddit traffic is from the U.S.

Image: <https://www.alexa.com/siteinfo/reddit.com>



New COVID cases per day in the U.S.

Image: Sawyer Click / [NBC News](#)

Reddit Activity Metrics

COVID-19 Discussion Rates



COVID-19 lexicon (adapted from [1]) to compute daily % of posts that mention COVID-19

User Counts

Daily # of unique users who post

Discussion Content

LIWC dictionary²: daily mean % of words in post in each category

LDA topic model³: daily mean posterior probability of topic across posts

[1] Xiaolei Huang, et al. 2020. Coronavirus Twitter Data:A collection of COVID-19 tweets with automated annotations.

[2] James W Pennebaker, Ryan L Boyd, Kayla Jordan, and Kate Blackburn. 2015. The development and psy-chometric properties of LIWC 2015.

[3] David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent Dirichlet Allocation. *Journal of Machine Learning Research*.

Family &
Home



Anxiety



School



Information
Sharing



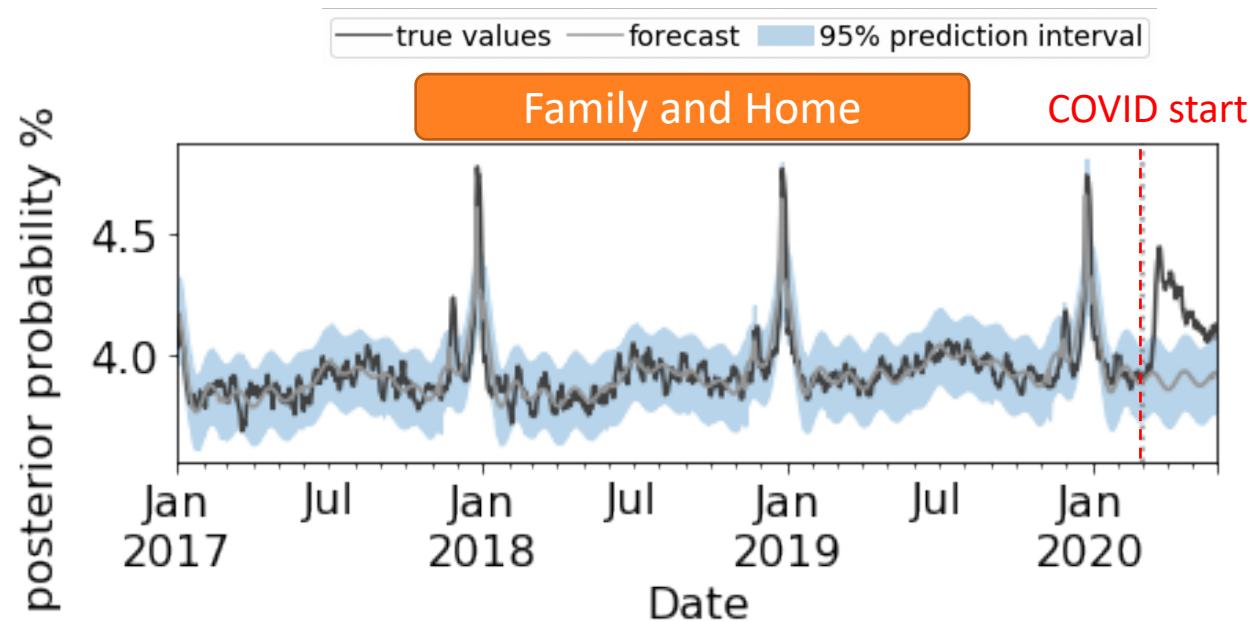
Topics

K=25 topics selected based on coherence scores

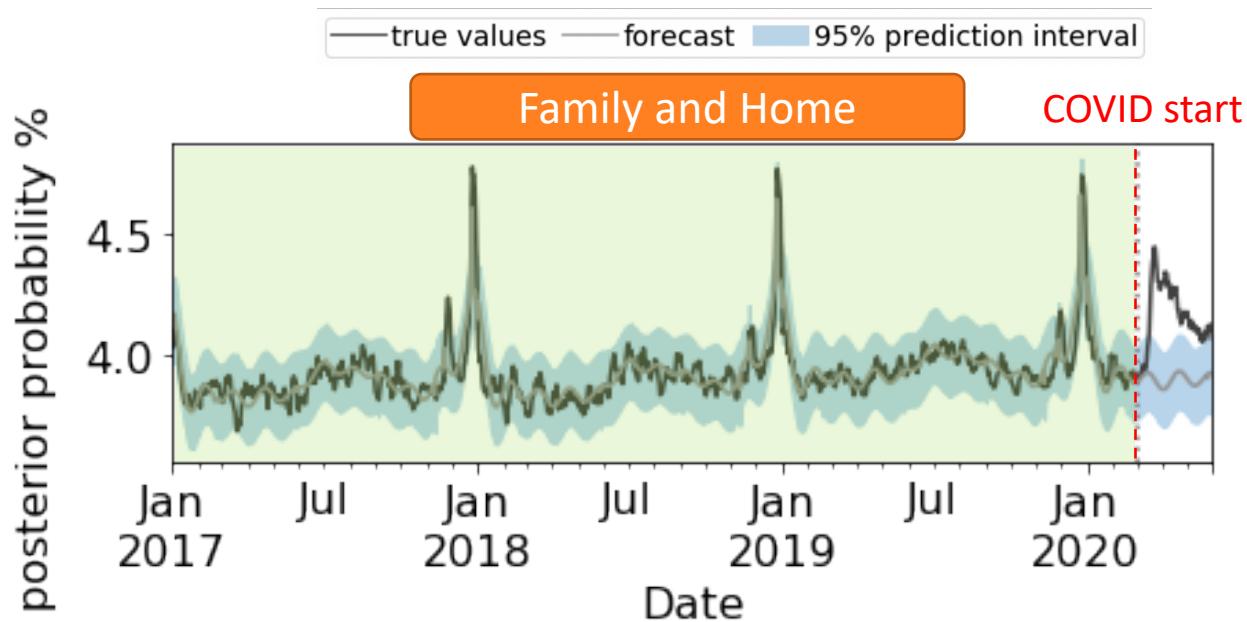
Themes include:

- **daily life concerns:** work, school, sleep and routine
- **personal relationships:** friends, family and children, relationships
- **mental health concerns:** anxiety, suicide, medical treatment

Time-Series Intervention Analysis

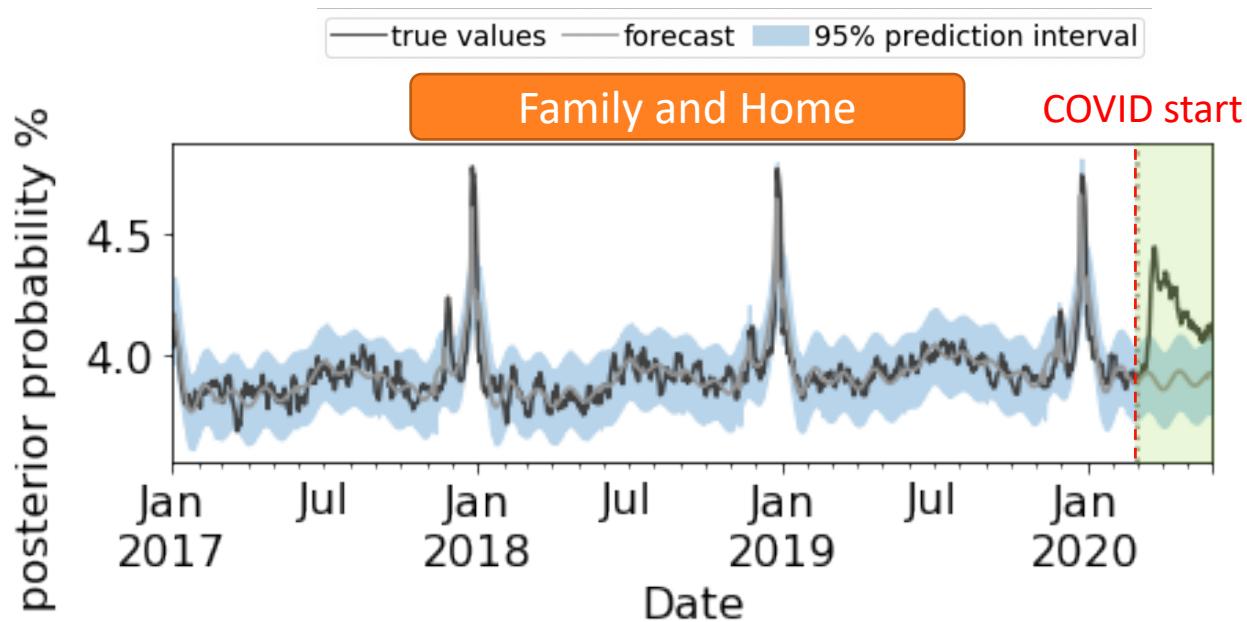


Time-Series Intervention Analysis



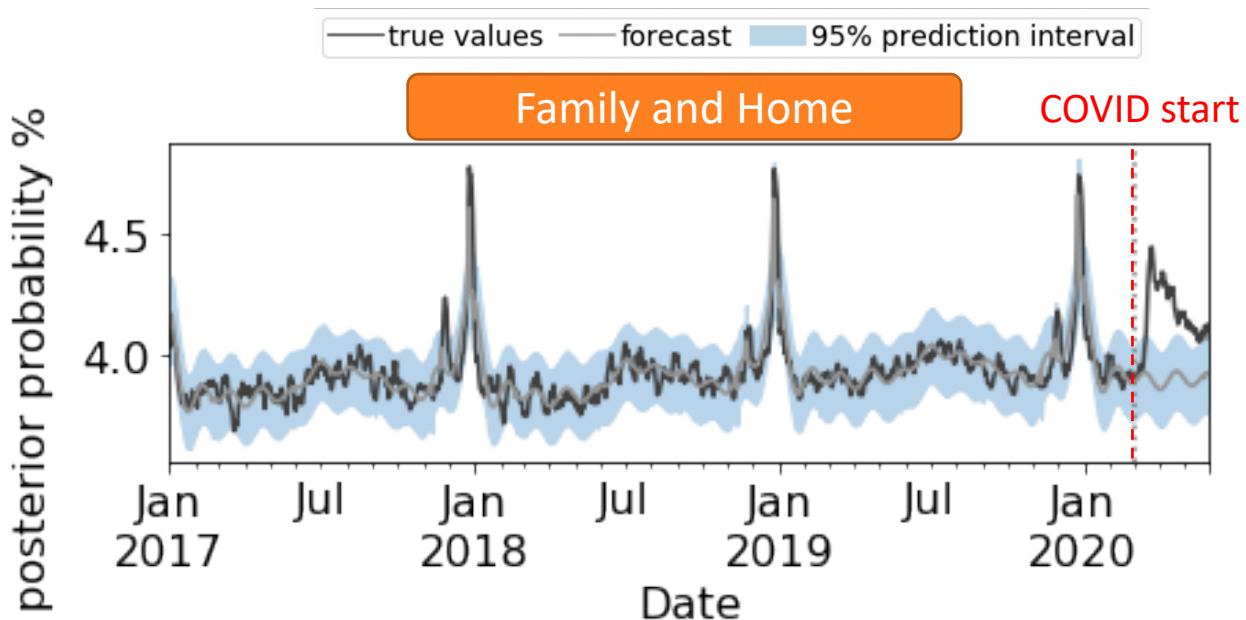
1. Fit time-series model to pre-COVID data

Time-Series Intervention Analysis



1. Fit time-series model to pre-COVID data
2. Forecast post-COVID values

Time-Series Intervention Analysis



1. Fit time-series model to pre-COVID data
2. Forecast post-COVID values
3. Examine difference between forecasted and observed values

Compute 95% prediction interval for post-COVID period

Null hypothesis (no change due to COVID-19): $\leq 5\%$ of observations outside of prediction interval

Apply **one-sample proportion test** to confirm/reject null hypothesis

Time-Series Model

Prophet model⁴ created by Facebook

Additive regression model with three components:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

trend approximated with
piecewise linear model

Time-Series Model

Prophet model⁴ created by Facebook

Additive regression model with three components:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

seasonality approximated with
Fourier series

Time-Series Model

Prophet model⁴ created by Facebook

Additive regression model with three components:

$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

holidays captured with
regression model

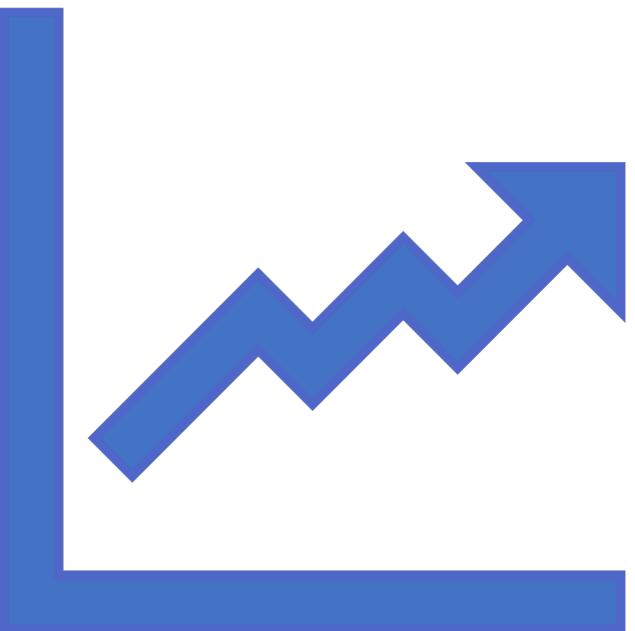
Time-Series Model

Prophet model⁴ created by Facebook

Additive regression model with three components:

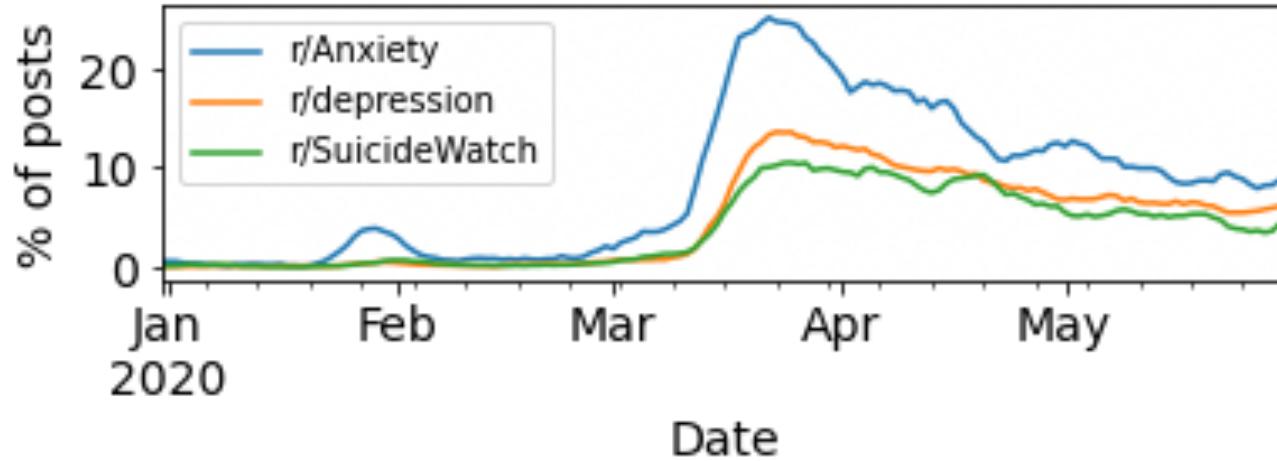
$$y(t) = g(t) + s(t) + h(t) + \epsilon_t$$

error



Results and Discussion

How often is COVID-19 discussed in mental health subreddits?

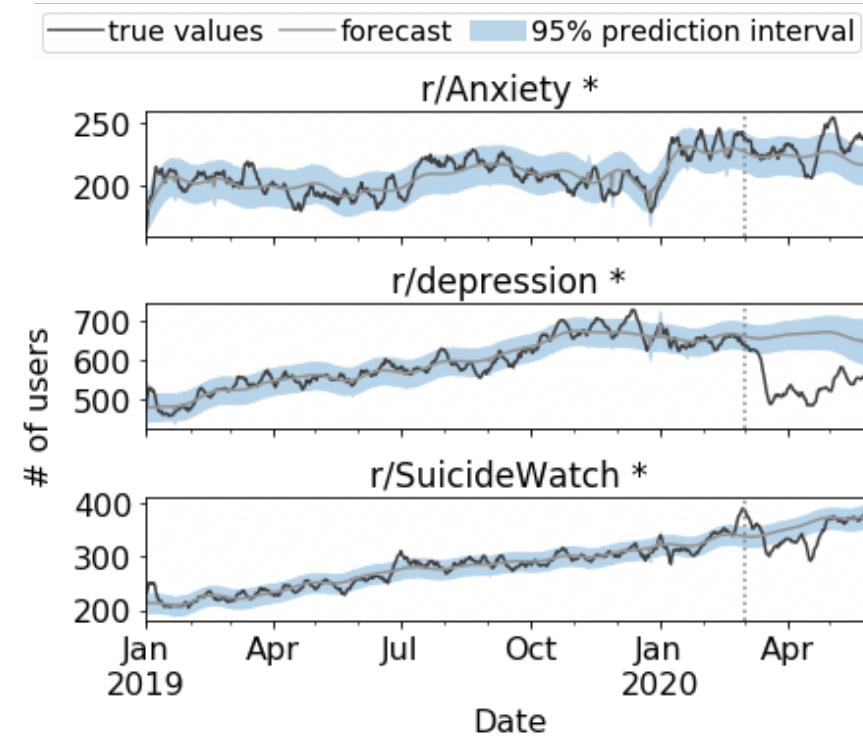


Percent of posts mentioning COVID-19 related words across
mental health subreddits

- Uptick in **March**, informing our choice of the post-COVID period
- COVID-19 is discussed **earlier** and **more often** on r/Anxiety

Has COVID-19 changed the number of users posting in mental health subreddits?

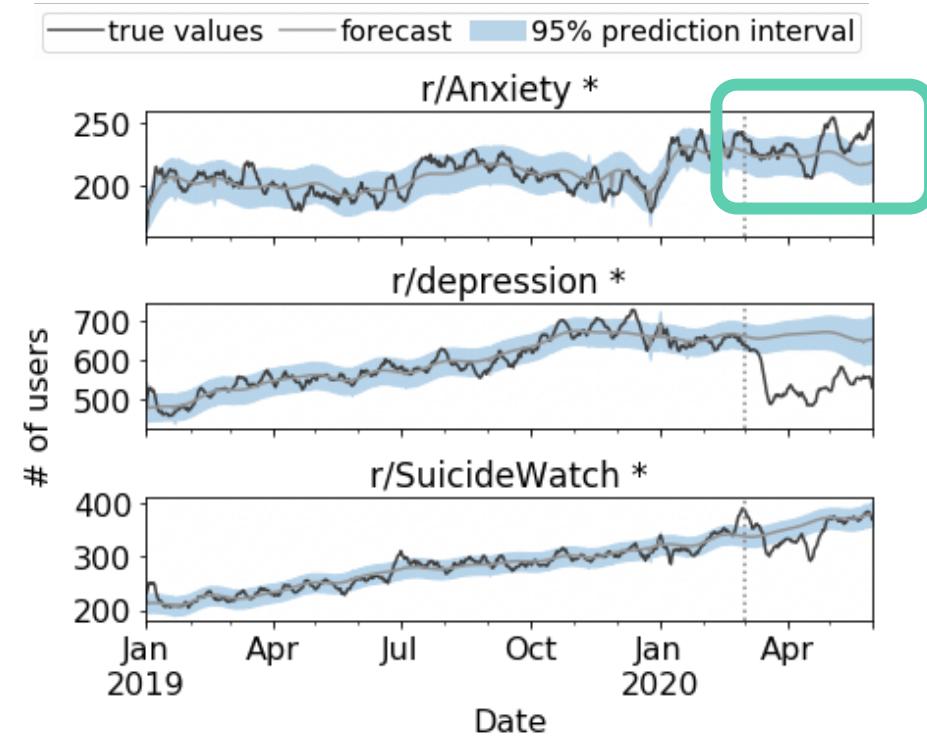
- Results from prophet modeling
- Intervention = March 2020
- * next to subreddit = statistically significant number of outliers



Daily active users over time

Has COVID-19 changed the number of users posting in mental health subreddits?

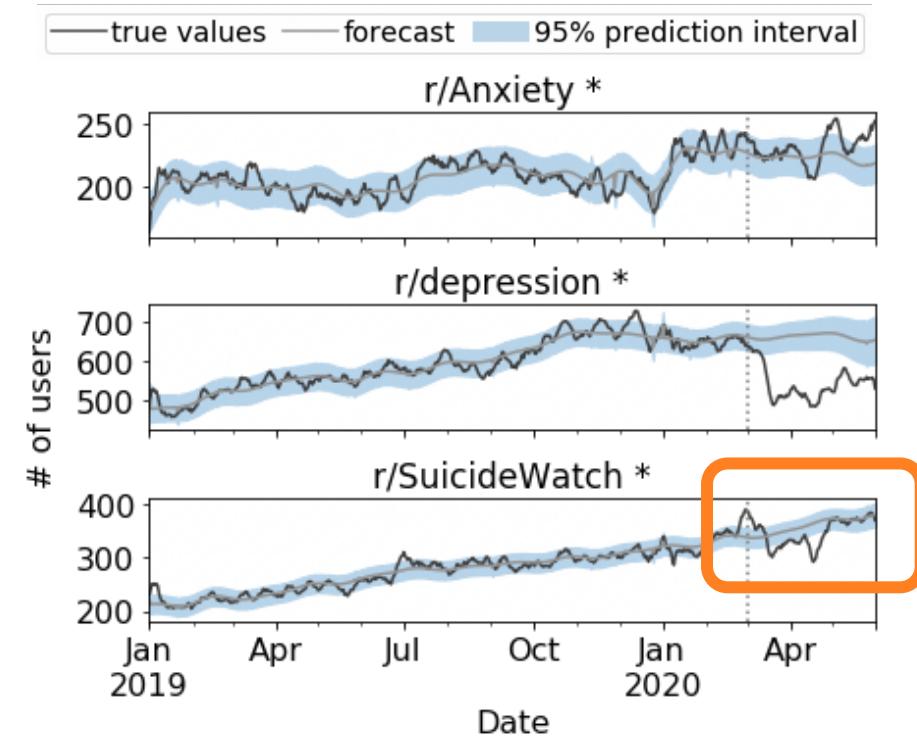
- There is a **small increase** on r/Anxiety



Daily active users over time

Has COVID-19 changed the number of users posting in mental health subreddits?

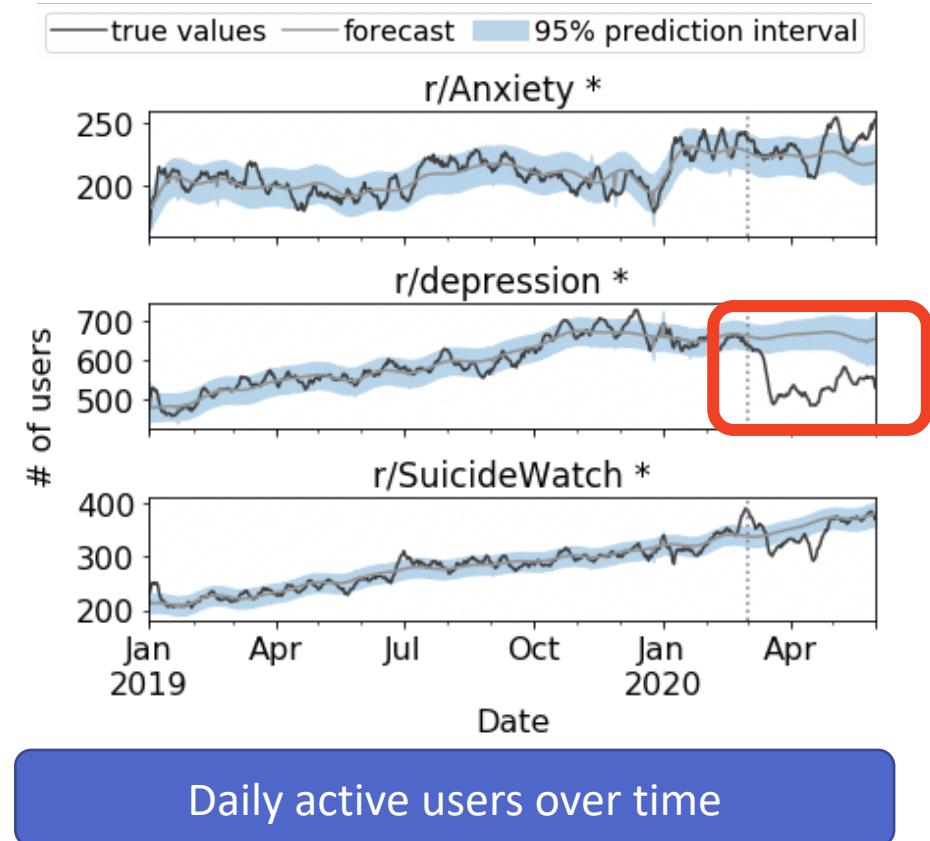
- There is a **small increase** on r/Anxiety
- There is a **small decrease** on r/SuicideWatch in mid-March – April, but activity level eventually becomes normal



Daily active users over time

Has COVID-19 changed the number of users posting in mental health subreddits?

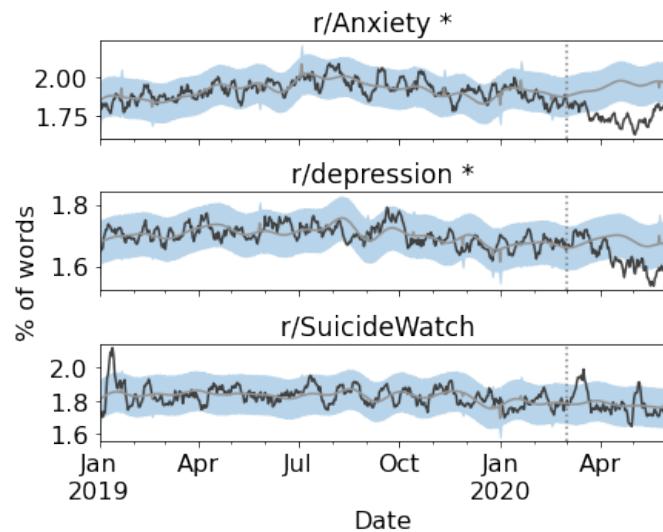
- There is a **small increase** on r/Anxiety
- There is a **small decrease** on r/SuicideWatch in mid-March – April, but activity level eventually becomes normal
- There is a **clear decrease** on r/depression that persists through the end of May



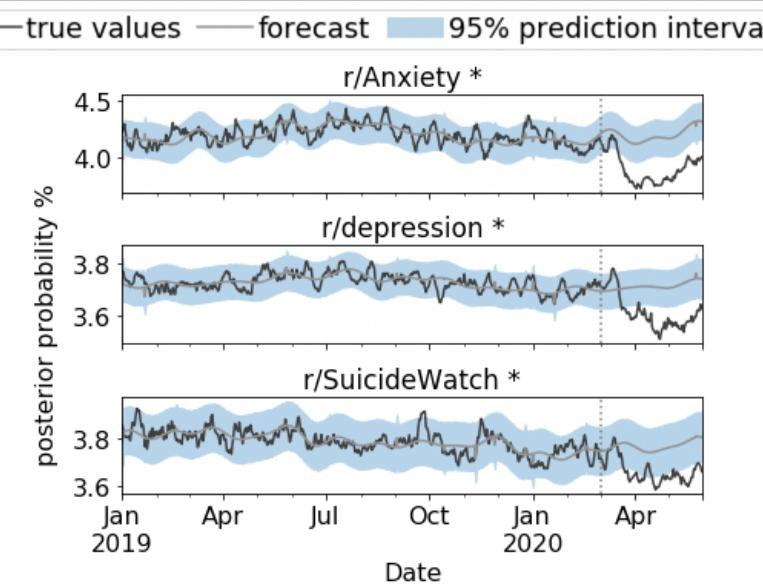
Has COVID-19 led to changes in
the discussions users have
surrounding mental health?

We use two types of metrics: **LIWC categories** and **LDA topics**

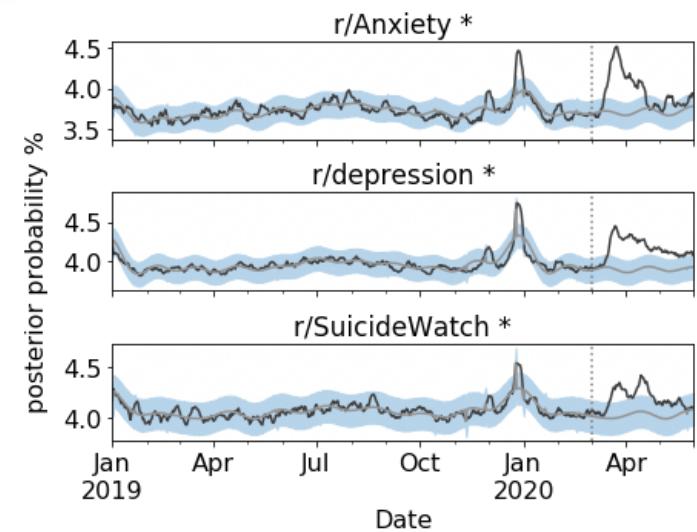
The Pandemic and Daily Life



Motion

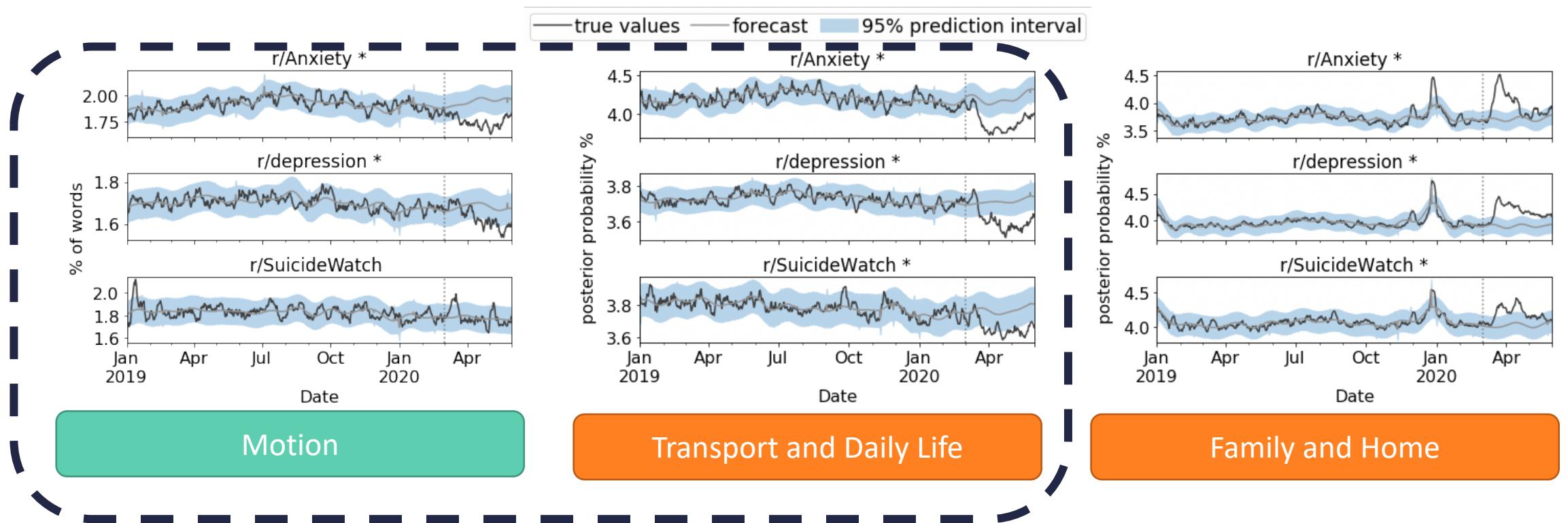


Transport and Daily Life

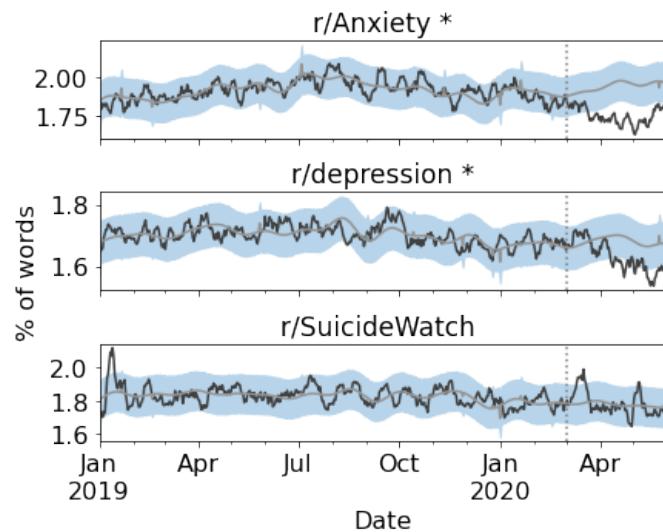


Family and Home

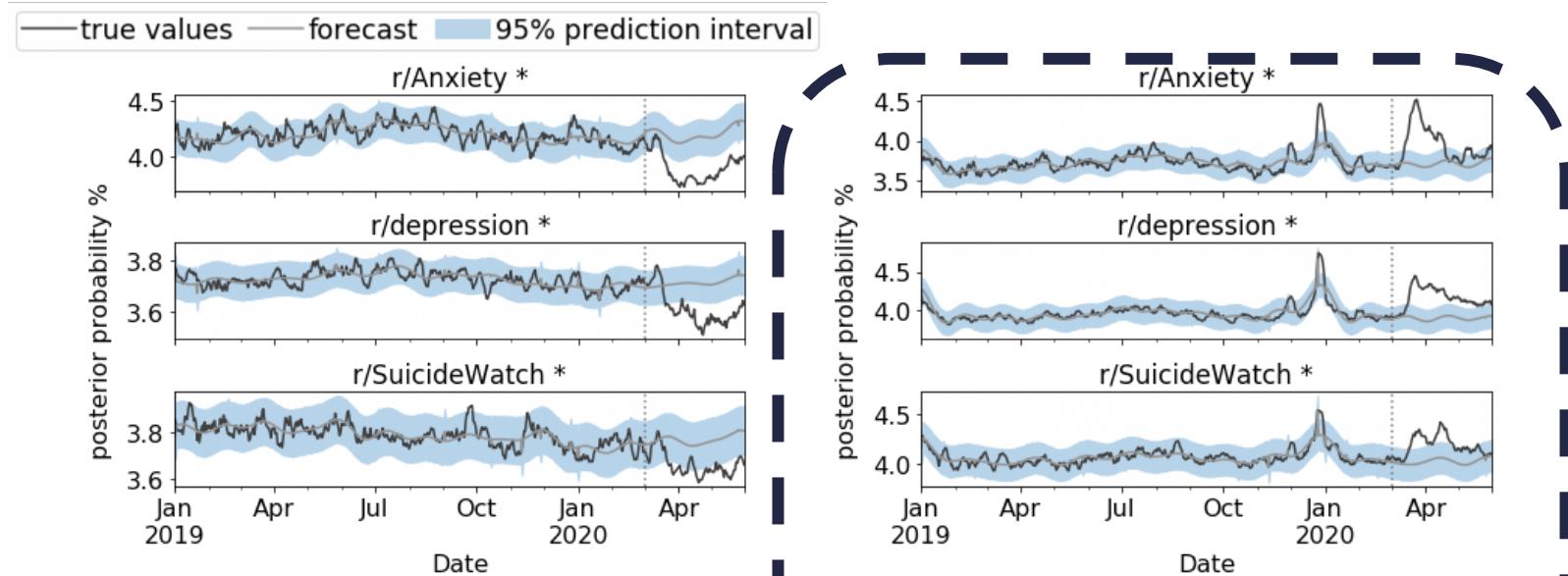
The Pandemic and Daily Life



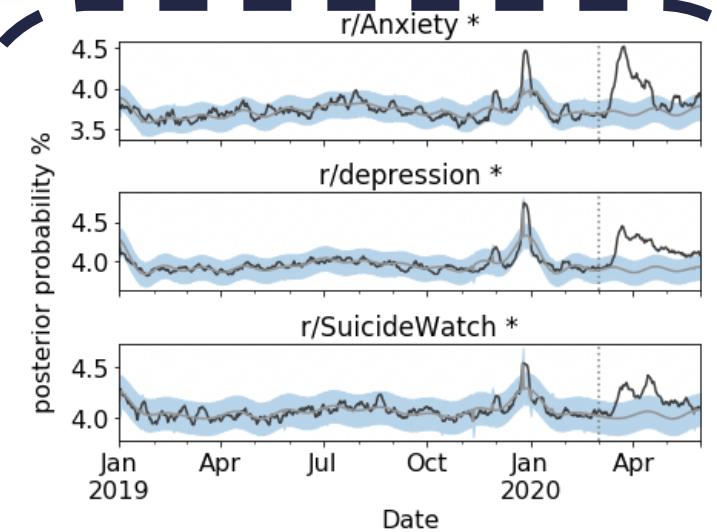
The Pandemic and Daily Life



Motion

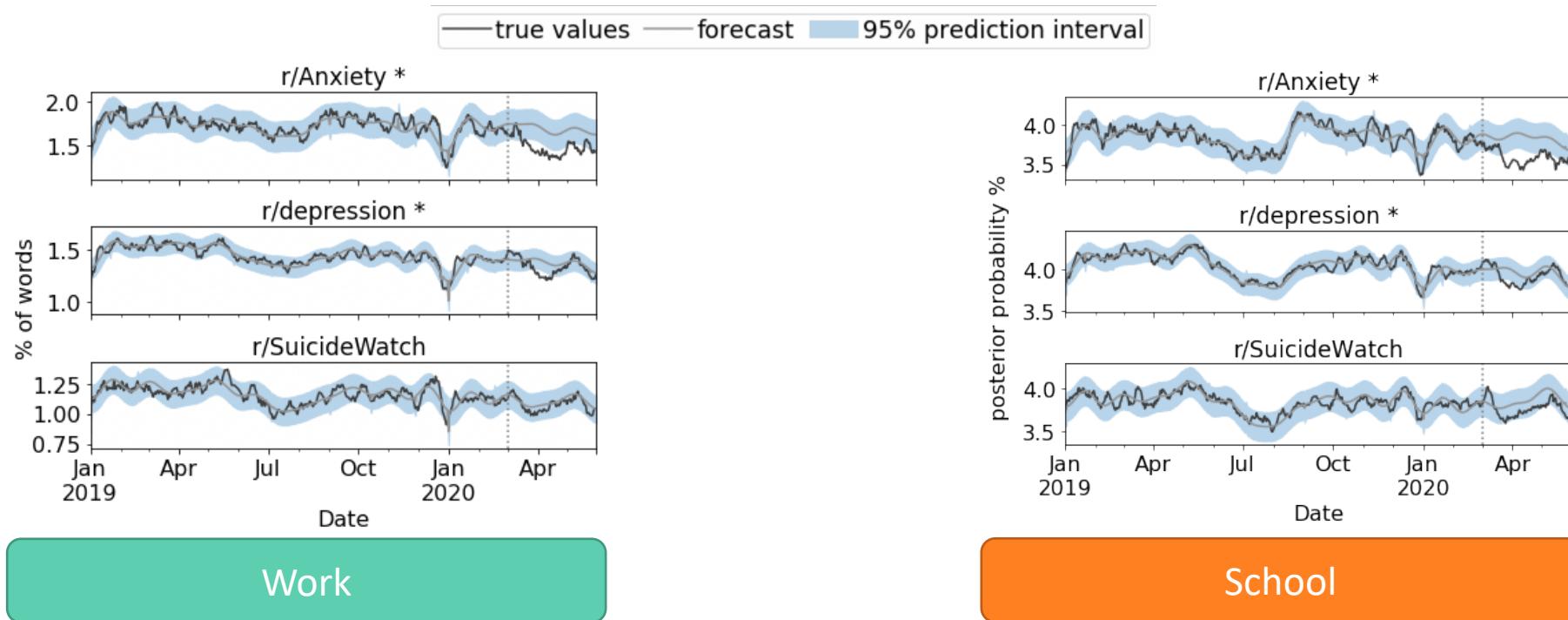


Transport and Daily Life

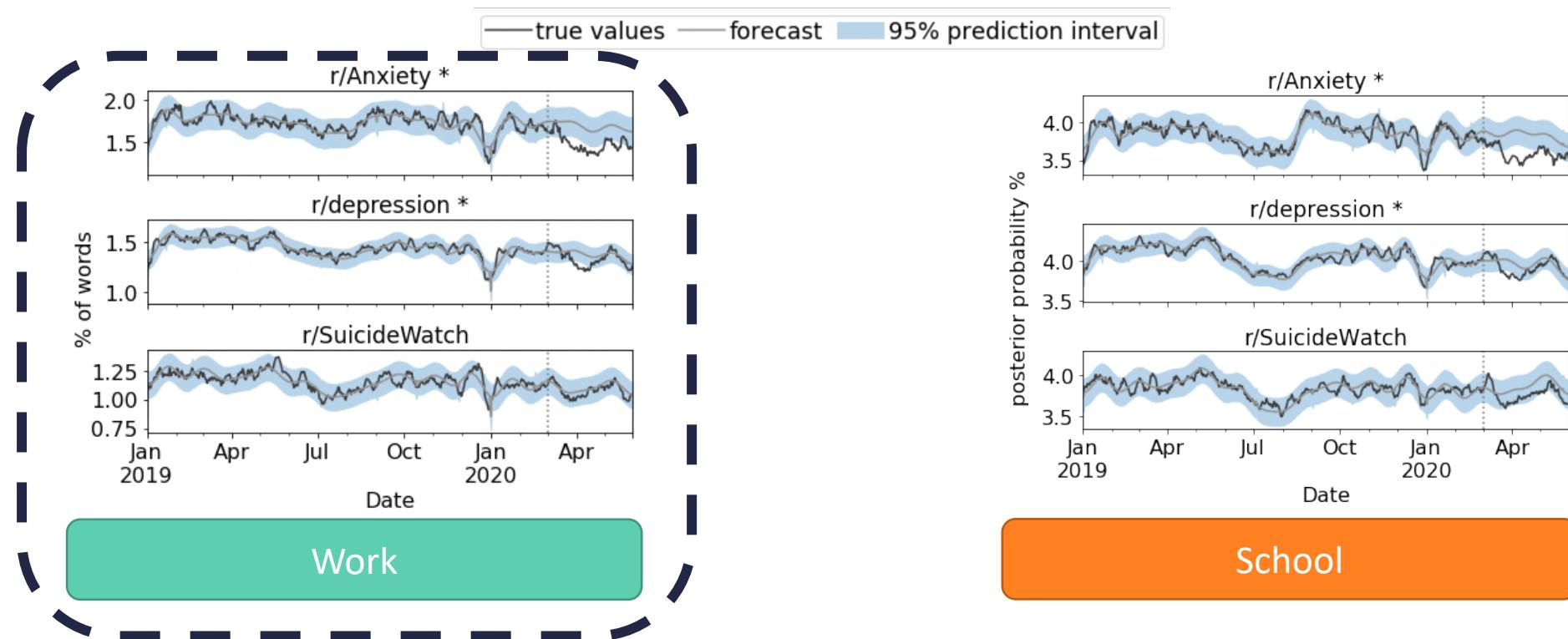


Family and Home

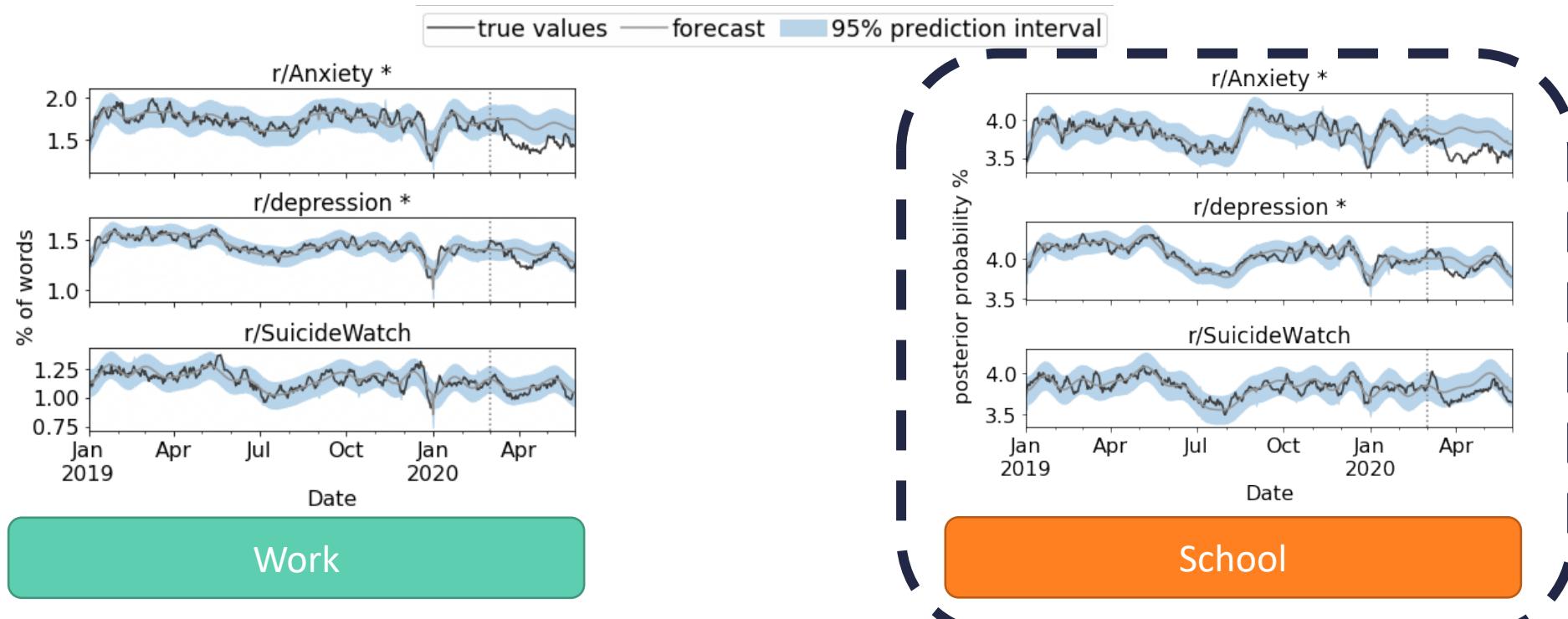
Work and School



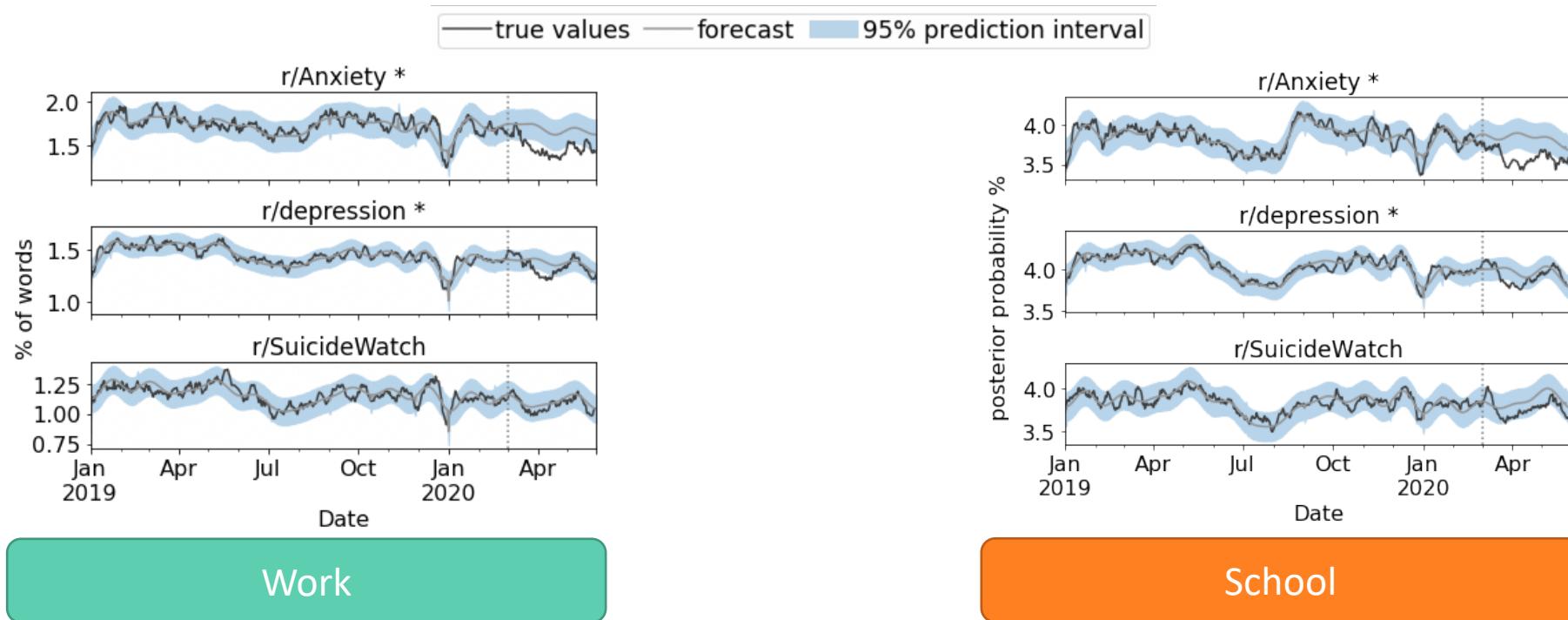
Work and School



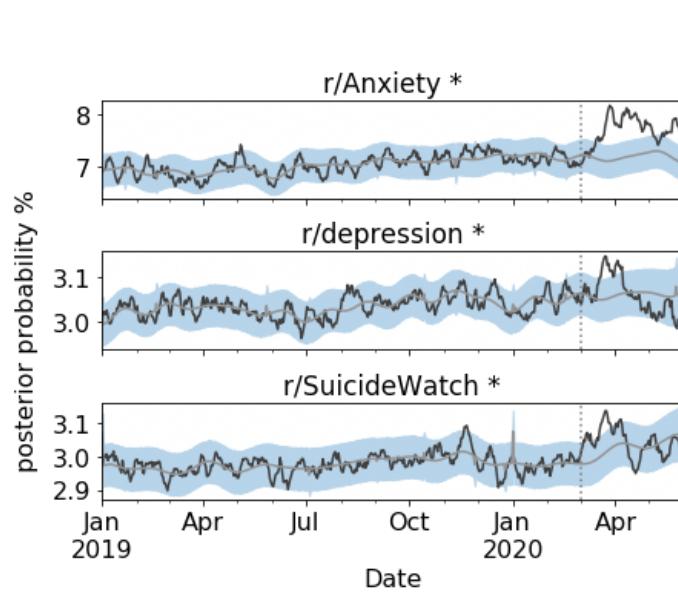
Work and School



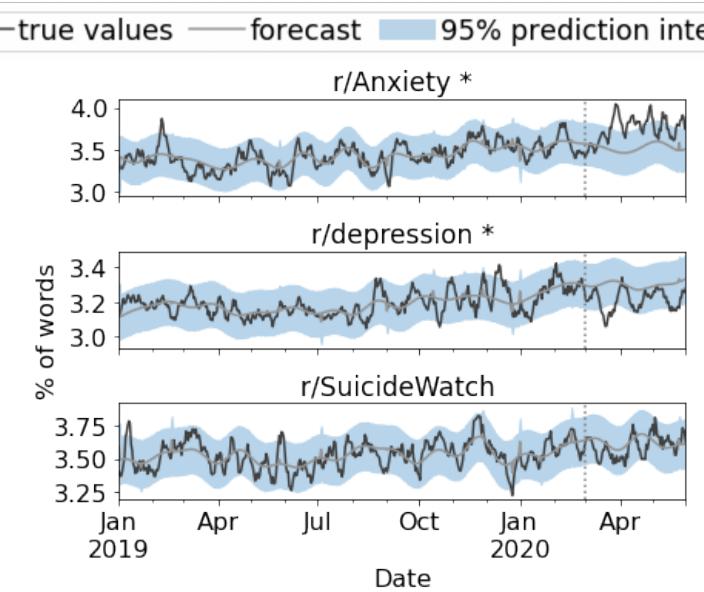
Work and School



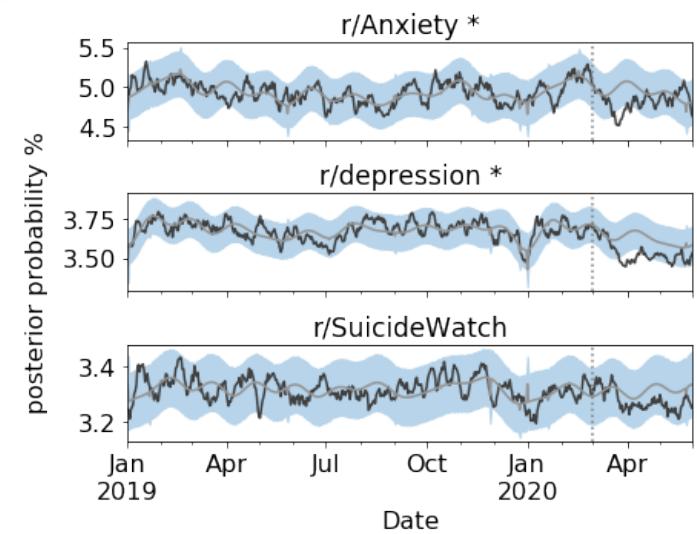
Anxiety and Health



Anxiety

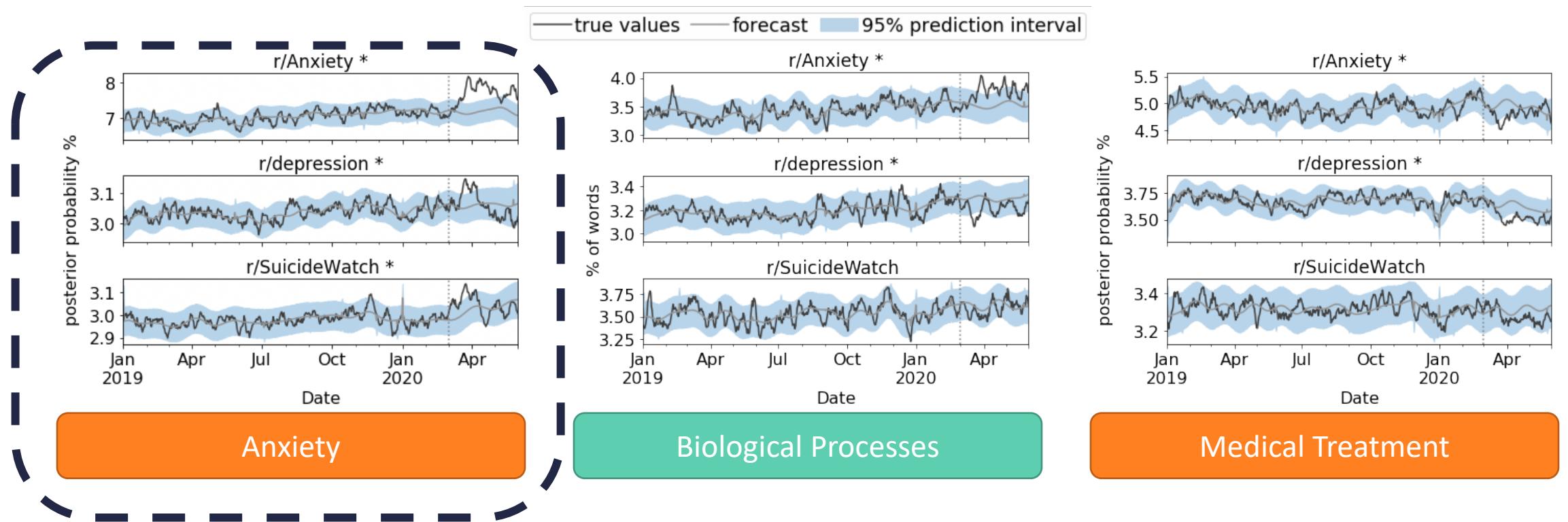


Biological Processes

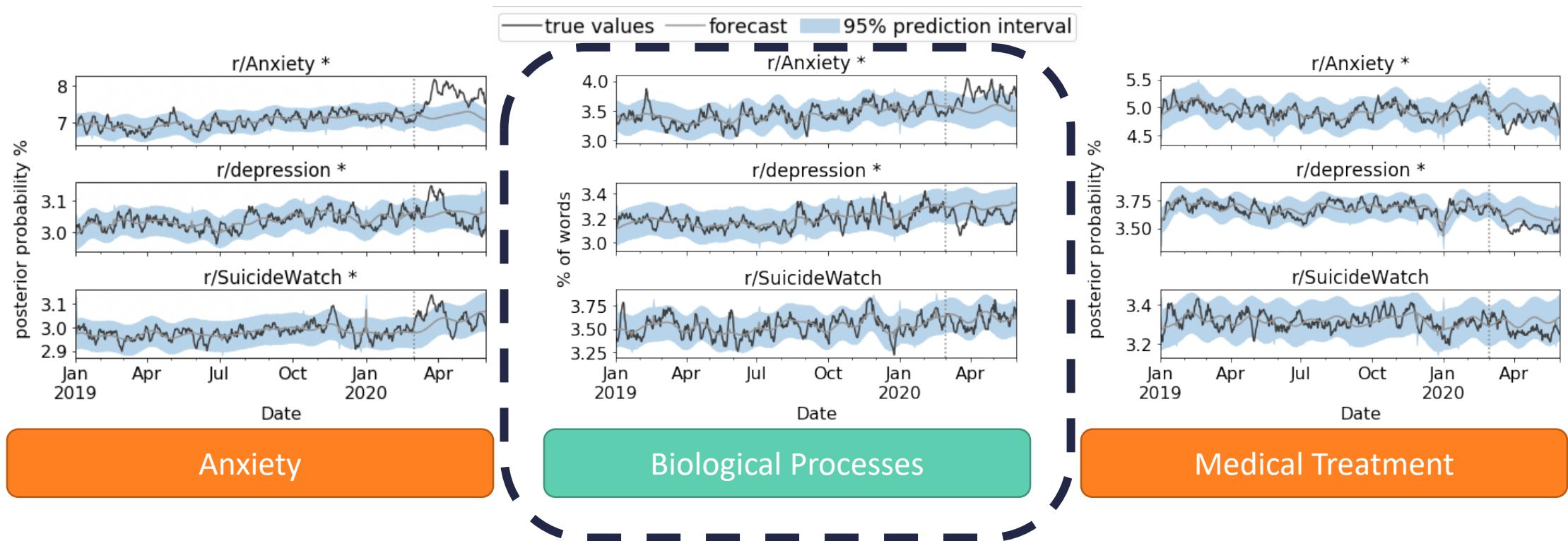


Medical Treatment

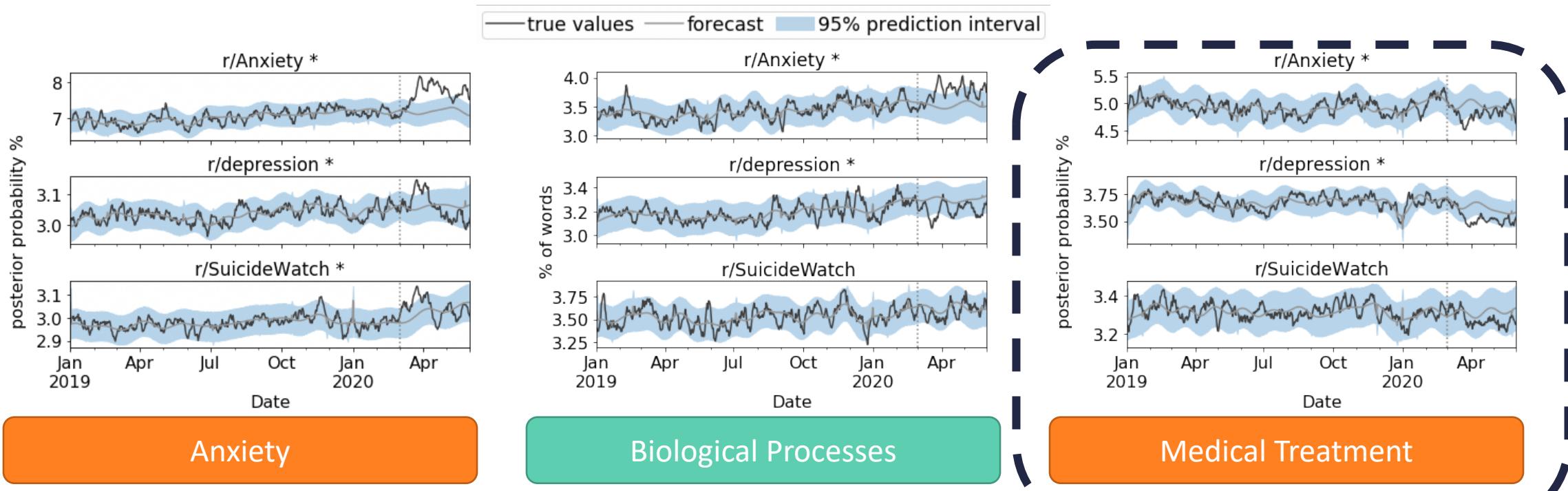
Anxiety and Health



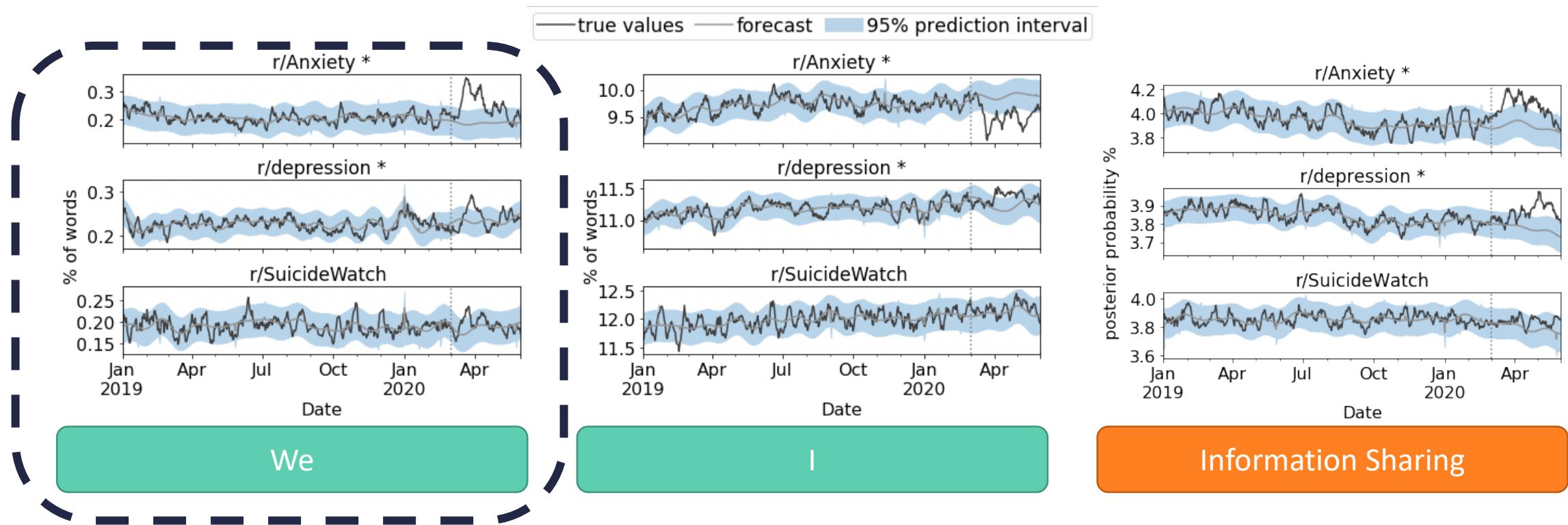
Anxiety and Health



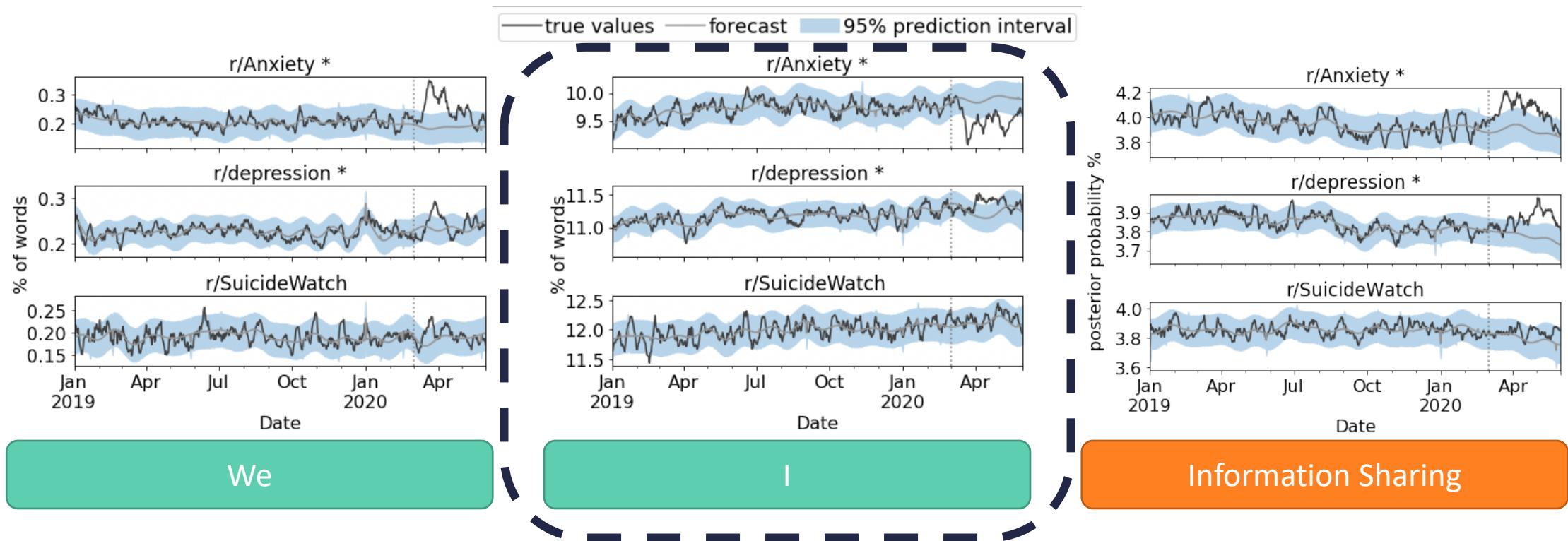
Anxiety and Health



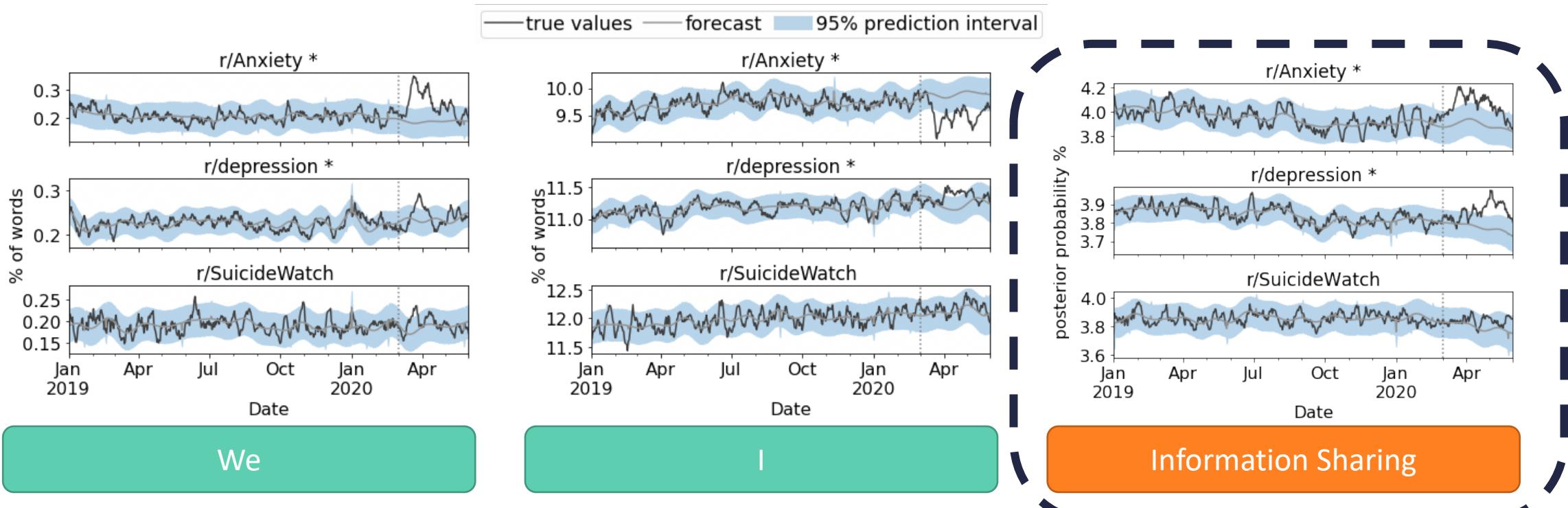
Collective Experience and Coping



Collective Experience and Coping



Collective Experience and Coping



Conclusions

- COVID-19 has had a **clear impact** on mental health subreddits
- The virus is **discussed most frequently** on **r/Anxiety**, and the subreddit also has an increase in activity
- Content has shifted, including in some **negative ways**: increase in discussion of anxiety and decrease in medical treatment
- However, there is **room for hope**, with increased information sharing and third-person plural usage indicating **collective coping**