Predicting COVID-19 outbreaks with Twitter data

Arran J. Davis, Cole B. J. Robertson, James Carney, Seán Roberts, and Dermot Lynott

Contents

1 Introduction **1**

1.1 Study hypotheses 1

2 Methods **3**

2.1 Tweet collection 3

3 References **3**

1 Introduction

Twitter data is increasingly being used to predict public health outcomes. Previous has shown that mentions of illnesses and illness symptoms by Twitter users can be used to predict future disease levels during epidemics and pandemics (Lopreite, Panzarasa, Puliga, & Riccaboni, 2021). Other research has used Twitter data to estimate regional variations in future orientation, showing that these regional variations predict risk-taking behaviours and disease outcomes (Ireland, Schwartz, Chen, Ungar, & Albarracín, 2015).

The studies presented here work to extend this research . . .

1.1 Study hypotheses

1.1.1 Study 1: Predicting COVID-19 outbreaks using a 'symptom talk' classifier on tweets from the UK in early 2020

This study would test two hypotheses related to changes in the frequency of tweets from the UK about illness symptoms (identified by a 'symptom talk' machine learning classifier).

1.1.1.1 Hypothesis 1.1

The frequency of tweets about illness symptoms is higher from 1 December 2019 - 1 April 2020 (early pandemic) than it is in pre-pandemic years during the same months. These tests could be as simple as t-tests, although they should probably use multilevel regression models that account for regional variations in symptom talk more generally.

Data: Tweet symptom talk frequencies from December 2019 - April 2020 would be compared to tweet symptom talk frequencies from the three previous previous years for the same months.

1.1.1.2 Hypothesis 1.2

UK locations that were particularly hard hit by COVID-19 in March and April 2020 (first wave in UK) would show greater pre-pandemic to early pandemic changes in the frequency of tweets about illness symptoms. These analyses could use a form of regression to test whether changes in the frequency of tweets about illness symptoms predicted COVID-19 outcomes at the regional level. Thus, change in symptom talk frequencies

(pre-pandemic years to pandemic years) for each region would be the predictor variable, and the outcome would be the COVID-19 cases or hospitalisations.

Data: Tweet symptom talk frequencies from the ~~flu~~ seasons of previous years (1 December to 1 April) would be compared to the same time period at the beginning of the pandemic.

1.1.1.3 Hypothesis 1.3

Tweet symptom talk frequencies can be used to predict changes in COVID-19 case frequencies over time. This could be a form of multilevel regression analysis that uses the frequency (at the regional level) of symptom talk in tweets at time ~~T1~~ to predict COVID-19 case loads or hospitalisation rates (at the regional level) at time T1 + ~14 days. ~~Analyses would thus take the form of a time series analysis.~~

Data: Tweet symptom talk frequencies from the first case in the UK (31 January 2020) to a year after the WHO officially declared a pandemic on 11 March 2020 (i.e., 31 January 2020 - 1 April 2021). Although 11 March 2021 marked a year since the WHO fficially declared a pandemic, 1 April 2021 would likely be a better end date for these analyses,

1.1.1.4 Considerations

- Sensorimotor norms will be used to give descriptive statistics for tweets identified as 'symptom talk' - this will be done to give readers a better understanding of the types of language being classified by the symptom talk classifier. Sensorimotor norms could be compared across time to identify changes in the types of symptoms that are being tweeted about.
- The symptom talk classider should be trained on *pre-pandemic* tweets.
- The frequency of illness talk should probably be at the individual level; the percentage of individuals using 'symptom talk' for a given time period (although overall frequency - ignoring users - may better estimate the severity of symptoms).
- General trends in symptom talk from flu season to flu season should also be examined to test whether changes in symptom talk are a result of broader historical trends (it could be that as time goes on people are generally more likely to use symptom talk on Twitter).
- Lopreite et al. (2021) cut-off their tweet sample on 21 January 2020, when the WHO announced COVID-19 and its symptoms, because there "would be no obvious way to disambiguate messages concerned with genuine local cases of pneumonia from messages elicited by mass media coverage of the outbreak" (p. 2). The symptom talk classifier somewhat avoids this problem, since it would be trained to identify when people are talking about any kind of symptom, but it is unclear to me what the correct methodology with regard to the date ranges. On the one hand, we would ideally use tweets from before media coverage affected peoples' behaviour and tweeting (i.e., those from January 2020 and earlier). On the other hand, focusing on tweets from January 2020 and earlier gives very little outcome data (COVID-19 infections) to work with, since cases in the UK are less than 10/day for all of February 2020.

1.1.2 Study 2: Predicting COVID-19 outbreak 'hot spots' using regional variations in Twitter users' time perspectives

This study would test a hypothesis related to the relationship between time perspectives (i.e., future orientation) and COVID-19 outcomes; future-orientation has been shown to predict a range of outcomes related to risky behaviours (e.g., Ireland et al., 2015).

~~1.1.2.1 Hypothesis 2.1~~

UK regions ~~with shorter time perspectives (fewer future-oriented tweets, etc.) will have worse COVID-19 outbreaks.~~

Data: Time perspectives will be assessed using the tweets from the three previous pre-pandemic flu seasons (i.e., the data made available from Study 1); when future orientation is measured should not exactly matter, as long as it is not *during* the pandemic.

1.1.2.2 Considerations

- It could also be interesting to see whether the ~~future orientation~~ of tweets changed from before to during the pandemic (potentially as a result of stress levels), although this is may be another study (which could use the data collected for both Study 1 and Study 2.

2 Methods

2.1 Tweet collection

The Twitter API (v2) for academic research was used to collect tweets for both studies. For all time periods in both studies, tweets were collected from all of the UK using the `place_country` field in the Twitter API query (Twitter, 2021). The UK country code (`GB`) was used in the `place_country` field; this returns all tweets with meta-data about tweet-specific locations in the UK (i.e., Twitter "Places" that are in the UK and specific GPS latitude/longitude point coordinates that can be associated with a Twitter Place that are in the UK).

2.1.1 Tweet collection dates

Tweet collection dates were meant to capture the 'three waves' of the COVID-19 pandemic in the United Kingdom (see Figure 1), as documented by Public Health England's data on COVID-19 cases and deaths (Public Health England, 2021). Tweet collection dates also included the three preceding (i.e., pre-pandemic) flu seasons.

Flu seasons were defined as lasting from 1 December in year `x` to 1 April in year `x + 1` - data suggest that case loads for influenza and other respiratory viruses generally peak in the UK during these four months (Public Health England, 2019). Using previous flu seasons as a comparison is crucial to analyses involving the symptom talk classifier, as this classifier was also likely measure influenza-related symptom talk; using pre-pandemic flu seasons as a benchmark helps to contorl for increases in symptom talk for reasons other than COVID-19 (e.g., influenza).

All tweets sent from the UK were thus collected for four seperate time periods:

- **1 November 2019 to 1 April 2021:** This time period encapsulates the first documented case of COVID-19 (thought to be in early November 2019) and the first year of the pandemic (the WHO officially declared a pandemic on 11 March 2020). The first case in the UK was recorded on 31 January 2020, and new cases exceed 1,000/day by the end of March 2020, and then peak (first wave) at over 5,000/day by mid April 2020. The second and third waves in the UK peak in November 2020 and around 1 January 2021, respectively.
- **1 December 2018 to 1 April 2019:** The most recent pre-pandemic flu season.
- **1 December 2017 to 1 April 2018:** The second most recent pre-pandemic flu season.
- **1 December 2016 to 1 April 2017:** The third most recent pre-pandemic flu season.

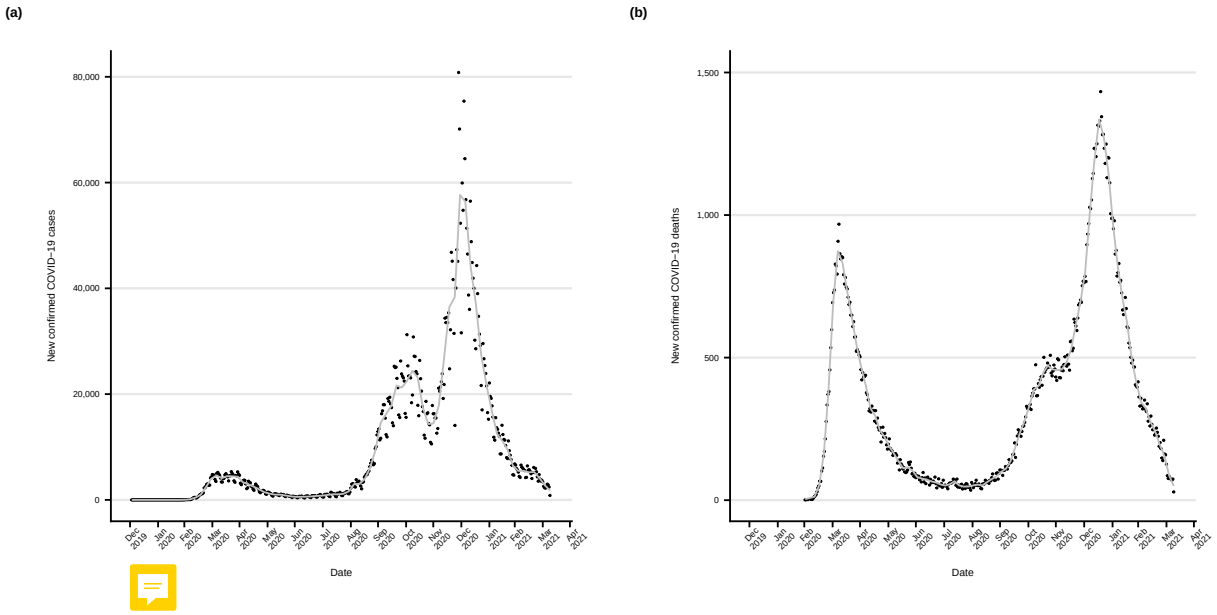

Figure 1: (a) Confirmed COVID-19 cases and (b) deaths attributed to COVID-19 in the United Kingdom from 1 December 2019 to one year after the WHO declared COVID-19 a pandemic (11 March 2021).

3 References

Ireland, M. E., Schwartz, H. A., Chen, Q., Ungar, L. H., & Albarracín, D. (2015). Future-oriented tweets predict lower county-level hiv prevalence in the united states. *Health Psychology*, *34*(S), 1252.

Lopreite, M., Panzarasa, P., Puliga, M., & Riccaboni, M. (2021). Early warnings of covid-19 outbreaks across europe from social media. *Scientific Reports*, *11*(1), 1–7.

Public Health England. (2019). Surveillance of influenza and other respiratory viruses in the uk: Winter 2018 to 2019. *Annual Flu Reports*.

Public Health England. (2021). *Coronavirus covid-19 in the uk*. Retrieved from https://api.coronavirus.data.gov.uk/v1/data

Twitter. (2021). *Filtering tweets by location*. Retrieved from https://developer.twitter.com/en/docs/tutorials/filtering-tweets-by-location