



COVID Symptom Tracker

Symptom Timing

Analysis by King's College London & ZOE

27 April 2020

Summary

1. We looked at the dataset for symptoms of healthy people who subsequently got sick.
2. Identifying potential COVID from a population of sick people calling NHS 111 is problematic, due to widely varying presentation. It is therefore well suited to an algorithm.
3. Fever or cough are not present in the first few days for most people with COVID.
4. A model optimised for track and trace should be different than the simple model used by King's as in press in *Nature Medicine*. It would likely use anosmia amongst other signals, and be optimised for symptoms across the first few days to maximise sensitivity. We are happy to build such a model if useful.

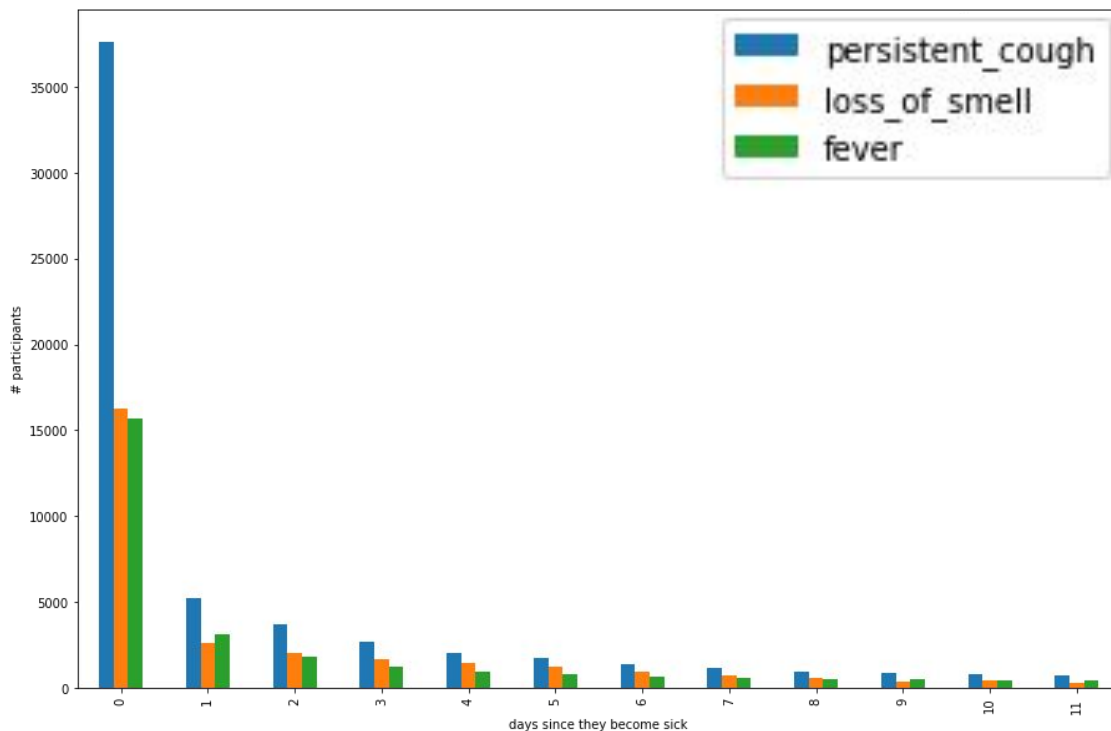
Symptom Timing:

Dataset Used

1. We start with 2.5M UK users logging their symptoms at home.
2. Then we select users who logged that they feel healthy for **at least the 3 first days of logging**, who then log that they **became sick**, and we then further reduce this to users who logged for **more than 10 days after they became sick**.
3. “App users who became sick throughout the study”. N = 116,568 users.
4. “App users who became sick throughout the study” and reported being tested and COVID positive (N = 727 users).

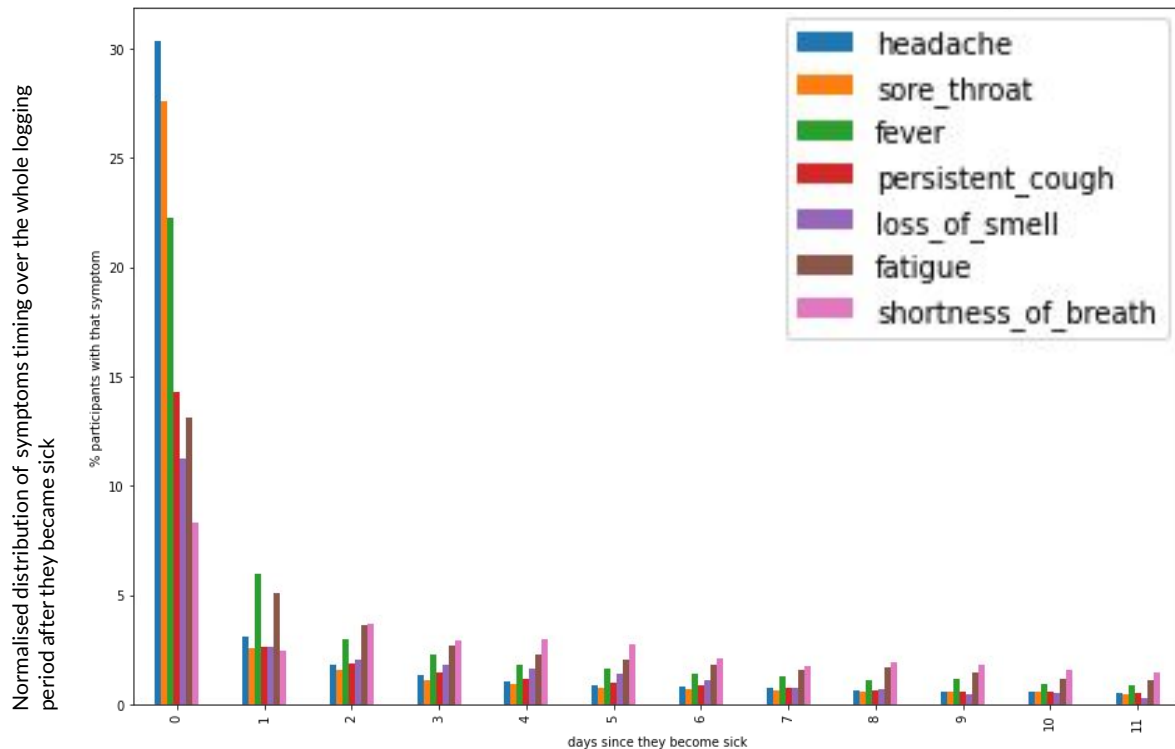
The first day that a user reports a specific symptom varies enormously across all users who became sick during the study.

Distribution of first day of symptom "X" over the whole logging period



- This data will contain all diseases, stress, etc - not just COVID. Just like any patient reporting to a GP or NHS 111.
- As many people logged loss of smell as fever on their first day of being sick.
- There's a broad variation in this distribution with some onset times happening later.

When we expand the symptom list we see that fatigue and shortness of breath are very visible later on, even in this population of everyone who is sick

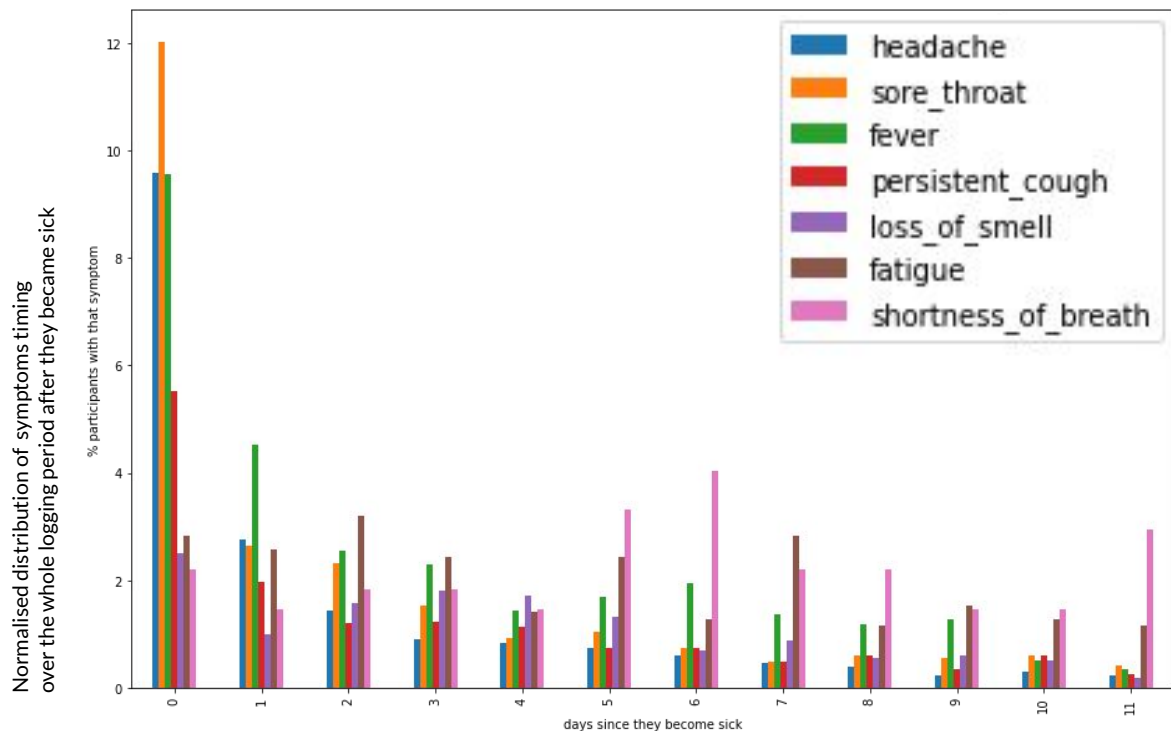


- While some symptoms are skewed towards the early days (headache, sore throat), others are more likely in the later stage (fatigue, shortness of breath)

Summary Data: App users who became sick throughout the study with full daily data (N=116,568)

	headache	sore_throat	fever	persistent_cough	loss_of_smell	fatigue	shortness_of_breath
count	55239.000000	44289.000000	6999.000000	14330.000000	6718.000000	4909.000000	1579.000000
mean	2.620739	2.764998	4.293185	4.026518	4.127717	5.976370	6.994934
std	5.327152	5.545912	6.373557	6.105831	5.918614	6.962925	6.841125
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	2.000000
50%	0.000000	0.000000	1.000000	1.000000	2.000000	3.000000	5.000000
75%	2.000000	2.000000	6.000000	6.000000	6.000000	9.000000	11.000000
max	30.000000	30.000000	29.000000	29.000000	30.000000	30.000000	30.000000

When we focus only on those tested for COVID-19 and were positive, the picture remains complex (N=727)



- This data shows that fever or cough are *not* present in the first few days for most people with COVID.
- The distributions of symptoms for users who tested positive are quite different from the overall population
- If we want to identify people for early testing, when viral shedding high, we should build a new model optimised for data in just the first few days.

Summary Data: App users who became sick and reported throughout the study and recalled tested COVID-19 +ve (N=727)

	headache	sore_throat	fever	persistent_cough	loss_of_smell	fatigue	shortness_of_breath
count	594.000000	426.000000	380.000000	481.000000	495.000000	232.000000	97.000000
mean	2.861953	3.232394	4.342105	4.467775	5.539394	6.612069	8.381443
std	4.688740	4.909533	5.204404	5.522256	5.423181	5.373933	5.754425
min	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	0.000000	0.000000	0.000000	0.000000	2.000000	2.000000	5.000000
50%	1.000000	1.000000	2.000000	2.000000	4.000000	5.500000	7.000000
75%	4.000000	5.000000	7.000000	7.000000	8.000000	10.000000	12.000000
max	29.000000	25.000000	23.000000	24.000000	27.000000	24.000000	27.000000