Self-reported symptoms as predictors of SARS-CoV-2 infection among people of the general population getting tested and living in the Amsterdam region, the Netherlands

5.1.2e

On behalf of the CONTROL team

ABSTRACT

Background

Since June 1, 2020, testing is available at GGD facilities to anyone with symptoms that may be indicative of a SARS-CoV-2 infection. For symptomatic people, immediate testing is advised and thus, most infections diagnosed at GGD test facilities are relatively early infections. Most SARS-CoV-2 symptoms are non-specific and are also common in other respiratory infections, leading to a large proportion of negative tests. In this analysis, we aimed to assess which symptoms best predict a positive test for SARS-CoV-2 in mostly symptomatic people of the general population who came for testing.

Methods

To study this, we used data from the SARS-CoV-2 test facilities of the Public Health Service of Amsterdam. Data were extracted from Coron-IT and anonymised before analysis. SARS-CoV-2 diagnosis was made by PCR. Our dataset contains 487,940 records of people tested from June 1st 2020 through December 31st 2020. We used logistic regression models with backward selection to obtain a model with only significant predictors ($p > 0.05$). We used a generalized estimating equation (GEE) model to correct for multiple tests of the same individual. The symptoms included as potential predictors are: coughing, fever, loss of smell, loss of taste, muscle ache, runny nose, shortness of breath, throat ache, or none of the above. Multivariable models were adjusted for age and gender. To assess whether the predictive value changed over time, we made separate models for each calendar month. Because of the large sample size and the resulting high power, there was the possibility that also very small and unimportant effects would come out statistically significant; therefore we also built a machine-learning-based predictive model using a random forest classifier. The outcome of this model is the importance of each variable, calculated using the mean decrease impurity (how much does the variable on average decrease the impurity of the split it makes).

Results

Overall, 51,599 of 487,940 tests were positive (10.6%). The median number of days between symptom onset and requesting a test among SARS-CoV-2 positive individuals was 2 (IQR:1-4). The most common symptoms were runny nose (58.1%), throat ache (50.2%), and cough (38.4%). These were common both in those with a positive and negative test. Loss of smell and loss of taste were uncommon in those that tested negative (5.0% and 4.9%), but more prevalent among the group that tested positive (15.6% and 15.5%). The mean number of symptoms reported per person was 2.0 (sd = 1.3). The prediction model showed that all symptoms contributed significantly to predicting a positive test result, but the strongest predictors were: cough (OR: 1.60, 95% CI: 1.57-1.63), fever (OR: 2.11, 95% CI: 2.07-2.16), loss of smell (OR: 2.21, 95% CI: 2.12-2.29), loss of taste (OR: 1.98, 95% CI: 1.90-2.05), and muscle ache (OR: 2.45, 95% CI: 2.39-2.51). The predictive value of muscle ache, loss of smell, and loss of taste seemed to decline slightly from June through December, whereas those of cough and fever slightly increased. In the random forest model results, the four features with the

highest feature importance are fever, loss of smell, loss of taste, and muscle ache, which implies that these four features impact the decision most on average. This corresponds with the findings of the logistic regression model. Furthermore, gender, 'none of the above' and age are the variables with the lowest importance, with an average of under 0.05.

Conclusions

Cough, fever, loss of smell or taste, and muscle ache appear to be the best predictors of a SARS-CoV-2 diagnosis in symptomatic people of the general population who were tested, whereas runny nose, shortness of breath, and throat ache are more indicative of a negative test result. The positive results for loss of smell and loss of taste were expected, based on earlier reports of these symptoms being quite specific for SARS-CoV-2 infections. However, the relatively low prevalence of these two specific symptoms amongst those with a positive test result indicates that the absence of these symptoms cannot rule out infection in the general population. The negative results for runny nose and throat ache are in line with expectations, because these are typical symptoms of other common infections. The lack of predictive value of shortness of breath should be interpreted in relation to timing; this symptom is not a good indicator of early infection, but might be highly predictive of SARS-CoV-2 infection later in the course of infection, for instance in hospital settings. It should be noted that the symptoms in these analyses were self-reported and bias might be present. Symptoms were collected either online through a short self-assessment were the client needs to report at least one symptom to proceed in the request process or symptoms were assessed by phone operators who may not have assessed all symptoms accurately due to time pressure. Both under- or overrepresentation of symptoms might be present.

Table 1: The prevalence (%) of symptoms in individuals by SARS-CoV-2 test result and overall, from June 2020 through December 2020, the Amsterdam region, the Netherlands

Symptom	Negative	Positive	Total
Coughing	37.2	48.2	38.4
Fever	15.4	32.1	17.2
Loss of smell	5.0	15.6	6.1
Loss of taste	4.9	15.5	6.0
Muscle ache	9.3	23.8	10.8
Runny nose	59.0	50.5	58.1
Shortness of breath	13.2	13.8	13.3
Throat ache	51.5	39.4	50.2
None of the above*	5.7	5.5	5.6

* Some individuals were tested due to other reasons than symptoms (e.g. contacts of cases) or might have had other symptoms than the once listed in Coron-IT.

Table 2: The number and percentage of positive SARS-CoV-2 test results by number of symptoms, month, gender, and age group, from June 2020 through December 2020, the Amsterdam region, the Netherlands

	Negative		Positive		Total
	N	%	N	%	N
TOTAL	**436,341**	**89.4%**	**51,599**	**10.6%**	**487,940**
Number of symptoms					
0	39,583	90.6%	4,129	9.4%	43,712
1	129,214	92.5%	10,494	7.5%	139,708
2	143,491	90.4%	15,249	9.6%	158,740
3	81,913	87.6%	11,648	12.4%	93,561
≥4	42,140	80.7%	10,079	19.3%	52,219
Month					
June	23,368	98.6%	329	1.4%	23,697
July	31,595	97.8%	723	2.2%	32,318
August	64,369	95.5%	3,056	4.5%	67,425
September	82,845	89.9%	9,300	10.1%	92,145
October	88,010	82.4%	18,808	17.6%	106,818
November	54,374	86.7%	8,353	13.3%	62,727
December	91,780	89.3%	11,030	10.7%	102,810
Gender					
Female	241,498	90.1%	26,410	9.9%	267,908
Male	194,843	88.6%	25,189	11.4%	220,032
Age group (years)					
0-14	24,622	89.9%	2,776	10.1%	27,398
15-24	77,857	87.9%	10,673	12.1%	88,530
25-34	126,970	91.2%	12,262	8.8%	139,232
35-44	48,166	91.8%	4,312	8.2%	52,478
45-54	66,951	89.4%	7,953	10.6%	74,904
55-64	65,094	86.8%	9,889	13.2%	74,983
65-74	19,596	88.5%	2,545	11.5%	22,141
75+	7,085	85.6%	1,189	14.4%	8,274

Table 3: Effect size and importance per predicting variable in the multivariable prediction models for a positive SARS-CoV-2 test outcome, from June 2020 through December 2020, the Amsterdam region, the Netherlands

	Model 1			Model 2
	OR	95% CI		Importance
Cough	1.60	1.57	1.63	0.083
Fever	2.11	2.07	2.16	0.145
Loss of smell	2.20	2.12	2.29	0.115
Loss of taste	1.98	1.90	2.06	0.114
Muscle ache	2.45	2.39	2.51	0.156
Runny nose	0.70	0.69	0.72	0.057
Shortness of breath	0.68	0.66	0.70	0.047
Throat ache	0.57	0.56	0.58	0.088
None of the above	0.93	0.89	0.97	0.030
Gender (male)	1.11	1.09	1.13	0.042

Model 1: Generalized estimating equation model, adjusted for age; all predictors were retained after backward selection, with $p < 0.05$ as threshold for inclusion. Model 2: Machine-learning-based predictive model using a random forest classifier. OR = odds ratio, 95% CI = 95% confidence interval.

Figure 1: Odds ratios and confidence interval per variable in the full multivariable prediction model per month, from June 2020 through December 2020, the Amsterdam region, the Netherlands

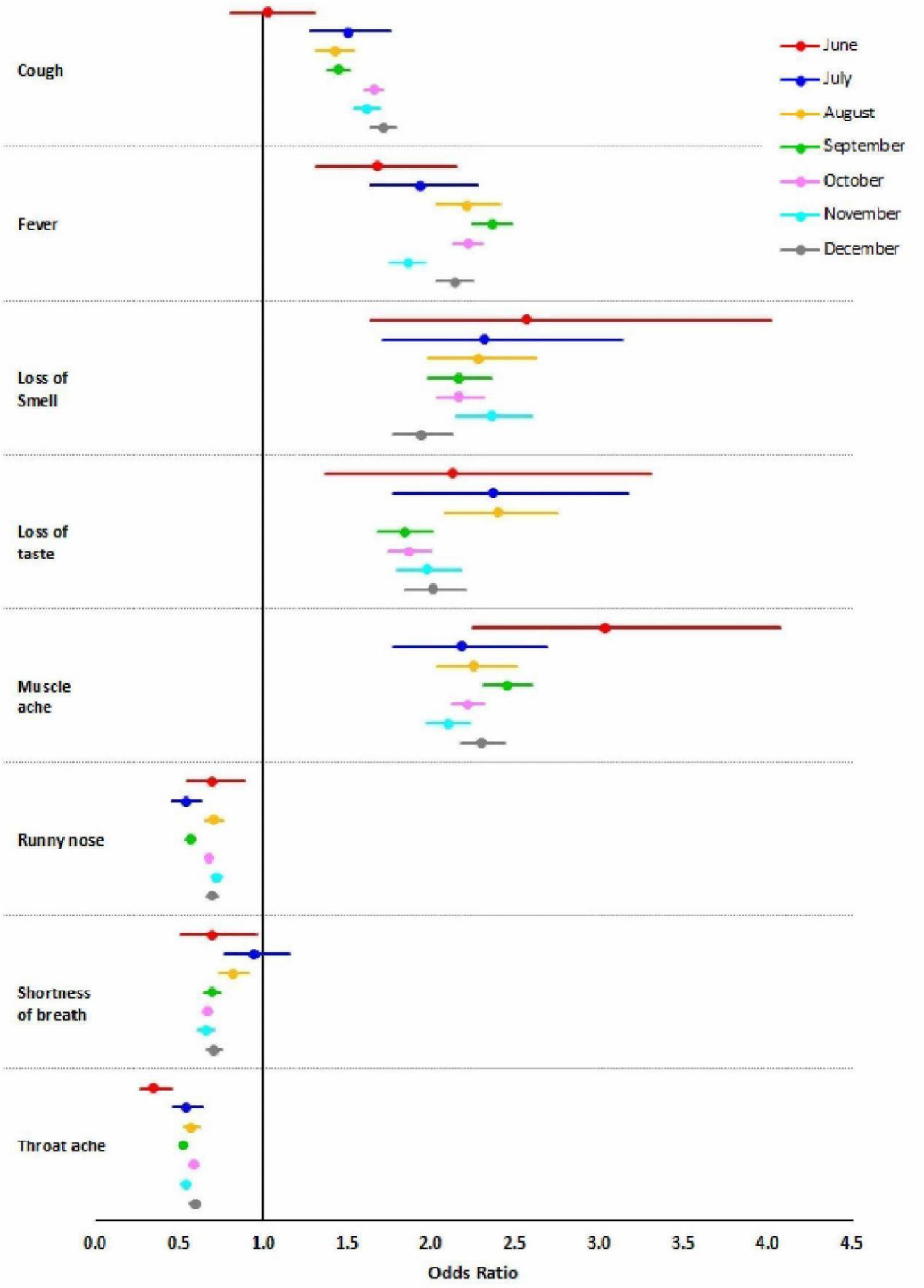

Table 4: Odds ratios and confidence per variable in the full multivariable prediction model per age group, from June 2020 through December 2020, the Amsterdam region, the Netherlands

	0-14 years			4-12 years *			15-24 years			25-34 years			35-44 years		
	OR	95% CI					OR	95% CI		OR	95% CI		OR	95% CI	
Cough	0.77	0.70	0.84	0.71	0.63	0.80	1.27	1.21	1.32	1.67	1.60	1.73	1.67	1.56	1.78
Fever	1.48	1.35	1.62	1.49	1.32	1.68	1.81	1.73	1.91	2.09	2.00	2.18	2.36	2.20	2.54
Loss of smell	1.75	1.42	2.15	1.66	1.20	2.30	2.15	1.99	2.31	2.30	2.14	2.46	2.87	2.54	3.24
Loss of taste	2.32	1.90	2.84	1.69	1.22	2.34	2.12	1.97	2.29	1.96	1.83	2.11	1.98	1.74	2.24
Muscle ache	2.16	1.82	2.58	2.04	1.57	2.66	1.98	1.86	2.11	2.24	2.14	2.36	2.69	2.48	2.91
Runny nose	0.64	0.59	0.70	0.69	0.61	0.77	0.79	0.76	0.83	0.79	0.76	0.82	0.69	0.65	0.74
Shortness of breath	0.74	0.62	0.88	0.69	0.53	0.88	0.74	0.69	0.79	0.75	0.71	0.80	0.62	0.56	0.69
Throat ache	0.61	0.56	0.67	0.56	0.49	0.63	0.63	0.61	0.66	0.60	0.58	0.63	0.49	0.46	0.53
None of the above	1.18	1.01	1.39	1.33	1.08	1.62	1.01	0.92	1.11	1.00	0.93	1.09	0.89	0.76	1.03
Gender	0.93	0.86	1.00	0.96	0.86	1.07	1.09	1.04	1.14	1.10	1.06	1.15	1.06	1.00	1.14
	45-54 years			55-64 years			65-74 years			75 years and above					
	OR	95% CI		OR	95% CI		OR	95% CI		OR	95% CI				
Cough	1.75	1.67	1.84	2.05	1.75	1.67	1.84	2.05	1.75	1.67	1.84	2.05			
Fever	2.41	2.28	2.55	2.36	2.41	2.28	2.55	2.36	2.41	2.28	2.55	2.36			
Loss of smell	2.51	2.27	2.77	1.85	2.51	2.27	2.77	1.85	2.51	2.27	2.77	1.85			
Loss of taste	1.86	1.68	2.05	1.78	1.86	1.68	2.05	1.78	1.86	1.68	2.05	1.78			
Muscle ache	2.90	2.74	3.08	2.69	2.90	2.74	3.08	2.69	2.90	2.74	3.08	2.69			
Runny nose	0.62	0.59	0.66	0.64	0.62	0.59	0.66	0.64	0.62	0.59	0.66	0.64			
Shortness of breath	0.61	0.57	0.66	0.61	0.61	0.57	0.66	0.61	0.61	0.57	0.66	0.61			
Throat ache	0.50	0.47	0.52	0.56	0.50	0.47	0.52	0.56	0.50	0.47	0.52	0.56			
None of the above	0.87	0.78	0.97	0.75	0.87	0.78	0.97	0.75	0.87	0.78	0.97	0.75			
Gender	1.02	0.97	1.08	1.24	1.02	0.97	1.08	1.24	1.02	0.97	1.08	1.24			

* This age group covers elementary-school-aged children