

Letters to the Editor

Is a COVID-19 prediction model based on symptom tracking through an app applicable in primary care?

Dagmar M. Haller^a, Paul Sebo^{a,○}, Benoit Tudrej^{b,c,*,○} and Hubert Maisonneuve^{a,b,c,○}

^aPrimary Care Unit, Faculty of Medicine, University of Geneva, Geneva, Switzerland, ^bUniversity College of General Medicine, University Claude Bernard Lyon 1, Lyon, France and ^cMermoz Primary Health Centre, Lyon, France

*Correspondence to Benoit Tudrej, Mermoz Primary Health Centre, 99 Avenue Jean Mermoz, 69008 Lyon, France; Email: dr.tudrej@gmail.com

‘We need a prediction score’. This sentence has most certainly been through the mind of most general practitioners (GPs) confronted with the challenge of identifying patients with potential COVID-19 in their practice. Just as the Centor/McIsaac risk stratification score for pharyngitis, a COVID-19 prediction score could help us triage patients, restricting the need for Reverse Transcription Polymerase Chain Reaction (RT-PCR) to patients in an intermediate risk category (1).

We thus commend Menni et al. for proposing a smart and ingenious method to predict potential COVID-19 using real-time symptom tracking through an app (2). We recognize the fantastic potential of their prediction score, combining loss of smell and taste, fatigue, cough and loss of appetite in prospectively identifying individuals at risk of having SARS-CoV-2 infection. However, as GPs at the forefront of early identification efforts in the community, we advise caution if applying this score in clinical practice. Both the

population involved in this study, and the nature of the prospectively collected real-time data from app users are potentially different from what can be expected in primary care.

We compared the main features of the population involved in Menni et al.’s study, and the performance of their score, with data from a cross-sectional study conducted between 24 March and 29 April 2020 in Lyon (France) involving nearly 1200 primary care patients undergoing RT-PCR testing for COVID-19 suspicion (Table 1) (3).

Applied to our data, in which the proportion of positive tests was 20%, the prediction model demonstrated poor calibration (Hosmer and Lemeshow $\chi^2 = 53.2$, P -value <0.001) and poor discrimination (area under the receiver operating characteristic curve ROC-AUC = 0.58 [95% Confidence Interval (CI) 0.56–0.61]). Applying a probability threshold of 0.5 (as proposed by Menni et al.), the sensitivity of the prediction model was 0.21 [95% CI 0.16–0.26], the specificity 0.96 [95% CI 0.95–0.98], the positive predictive value 0.59

Table 1. Characteristics of the study population and association between symptoms included in the score and positivity of the SARS-CoV-2 test ($n = 1177$)

	Tested positive for SARS-CoV-2	Tested negative for SARS-CoV-2	Adjusted P -value ^a
Number	239	938	
Female (%)	63.2	64.9	
Age (years)	47.3 (17.6)	46.7 (18.0)	
Answered questions on symptoms (n)	239	938	
Loss of smell and taste (%)	23.4	4.5	<0.001
Fatigue (%)	13.4	16.4	0.01
Cough (%)	51.9	49.7	0.49
Loss of appetite (%)	0.4	0.5	0.88

The results are presented as percentage values for dichotomous traits and as mean and standard deviation for age.

^aFor the Odds Ratio in SARS-CoV-2 positive compared to negative patients using multivariable logistic regression (adjusted for clustering within labs, gender and age group).

Table 2. Contingency table summarizing the data set ($n = 1177$)

SARS-CoV-2 test	No infection according to the prediction model	Infection according to the prediction model	Total
Negative	904	34	938
Positive	190	49	239
Total	1094	83	1177

[95% CI 0.48–0.70] and the negative predictive value 0.83 [95% CI 0.80–0.85]. These values can be calculated using the contingency table (Table 2) showing the number of true positives, false positives, true negatives and false positives. The score only modestly increased the pre-test probability of a positive SARS-CoV2 RT-PCR. And, in patients with a negative score, the risk of infection was still 17% (1094 individuals with a negative score, including 190 false negatives). In conclusion, applying this score in clinical practice would not sufficiently reduce the number of uninfected patients who are referred to RT-PCR testing, and it would lead to a large number of patients being misdiagnosed as not having COVID-19.

Why is the performance of the model applied to our patients inferior to that of Menni et al.'s study? In Menni et al., the patients using the app reported potential symptoms of COVID-19 and the result of SARS-CoV2 test in real time. These authors highlight that the self-report is a major limitation of their study. In our study, most patients were referred by their GP because they were complaining of COVID-like symptoms or else they were self-referred health professionals (3). As we have limited information about the reasons for testing in Menni et al.'s study, it is difficult to compare but it is likely that our population was a more symptomatic population due to the GPs' pre-testing triage. Indeed, nearly half the patients in our sample reported fever (45.4%), reflecting a common reason for GPs to refer patients to testing at the time the data were collected. This symptom was registered in the app by only a third of the patients undergoing RT-PCR testing. In addition, with the exception of loss of smell and taste, the symptoms included in the score are common unspecific symptoms for which patients present to primary care. Finally, the proportion of older patients and of male patients was higher in our study, more closely reflecting the usual primary care demography.

Real-time symptom collection through an app seems to be an attractive method to screen for potential COVID-19 and Menni et al.'s approach confirms the crucial value of specific symptoms, such as loss of smell and taste in the diagnosis of this infection (3–5). Yet the score they propose should not be applied as such for primary care patients as it does not appear to perform well in this population.

Declarations

Funding: the study was funded by departmental resources.

Ethics: the study was approved by the Ethics Committee of the Collège National des Généralistes Enseignants IRB0010804 on the 5 May 2020 (approval number: 200423163).

Conflict of interest: the authors declare no competing interests.

Data availability: data available on request.

References

1. Willis BH, Coomar D, Baragilly M. Comparison of Centor and McIsaac scores in primary care: a meta-analysis over multiple thresholds. *Br J Gen Pract* 2020; 70: e245–54.
2. Menni C, Valdes AM, Freidin MB *et al.* Real-time tracking of self-reported symptoms to predict potential COVID-19. *Nat Med* 2020. doi:10.1038/s41591-020-0916-2
3. Tudrej B, Sebo P, Lourdaux J *et al.* Self-reported loss of smell and taste in SARS-CoV-2 patients: primary care data to guide future early detection strategies. *J Gen Intern Med* 2020 9: 1–3. doi:10.1007/s11606-020-05933-9.
4. Giacomelli A, Pezzati L, Conti F *et al.* Self-reported olfactory and taste disorders in SARS-CoV-2 patients: a cross-sectional study. *Clin Infect Dis* 2020. doi:10.1093/cid/ciaa330
5. Benezit F, Le Turnier P, Declerck C *et al.* Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. *Lancet Infect Dis* 2020. doi:10.1016/S1473-3099(20)30297-8