COVID-Q: validation of the first COVID-19 questionnaire based on patient-rated symptom gravity

Spinato G^{1,2}, Fabbris C¹, Conte F³, Menegaldo A¹, Franz L¹, Gaudio P¹, Cinetto F⁴, Agostini C⁴, Costantini G³, Boscolo-Rizzo P⁵

1: Department of Neurosciences, Section of Otolaryngology and Regional Centre for Head and Neck Cancer, University of Padova, Treviso, Italy

2: Department of Surgery, Oncology and Gastroenterology, Section of Oncology and Immunology, University of Padova, Padova, Italy

3: Department of Psychology, University of Milano Bicocca, 20126 Milan, Italy

4: Department of Medicine, Clinical Immunology and Hematology, University of Padova, Treviso, Italy

5: Section of Otorhinolaryngology, Azienda Sanitaria Universitaria Integrata di Trieste, Trieste, Italy

Corresponding author:

Cristoforo Fabbris MD

Department of Neurosciences, University of Padova, Piazza Ospedale 1, 31100, Treviso, Italy - cristoforo.fabbris@gmail.com

Conflicts of Interest, Source of Funding and Compliance with Ethical Standards

The authors have no conflict of interest, funding or financial relationships. An informed consent has been obtained for any procedure involving the patients described in this article. The manuscript has not been submitted to more than one journal for simultaneous consideration. The manuscript has not been published previously (partly or in full). All of the authors have participated in the planning, writing or revising the manuscript.

ABSTRACT

Objectives

The aim of the present study was to develop and validate the CoronaVirus Disease – 2019 (COVID-19)

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice. Questionnaire (COVID-Q), a novel symptom questionnaire specific for COVID-19 patients, to provide a

comprehensive evaluation which may be helpful for physicians.. A secondary goal of the present study was to evaluate the performance of the COVID-Q in identifying subjects at higher risk of being tested positive for COVID-19.

Material and methods

Consecutive non-hospitalized adults who underwent nasopharyngeal and throat swab for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) detection at Treviso Hospital in March 2020, were enrolled. Subjects were divided into positive (cases) and negative (controls) in equal number. All of them gave consent and answered the COVID-Q. Patients not able to answer the COVID-Q due to clinical conditions were excluded.

Parallel Analysis and Principal Component Analysis were used to identify clusters of items measuring the same dimension. The Item Response Theory (IRT)-based analyses evaluated the functioning of item categories, the presence of clusters of local dependence among items, item fit within the model and model fit to the data.

Results

Answers obtained from 230 COVID-19 cases (113 males, and 117 females; mean age 55 years, range 20-99 years) and 230 controls (61 males, and 169 females; mean age 46 years, range 21-89) were analyzed. Parallel analysis led to the extraction of six components, which corresponded to as many clinical presentation patterns: asthenia, influenza-like symptoms, ear and nose symptoms, breathing issues, throat symptoms, and anosmia/ageusia. The final IRT models retained 27 items as significant for symptom assessment. The total score on the questionnaire was significantly associated with positivity to the molecular SARS-CoV-2 test: subjects with multiple symptoms were significantly more likely to be affected by COVID-19 ($p < .001$). Older age and male gender also represented risk factors. Presence of breathing issues and anosmia/ageusia were significantly related to positivity to SARS-CoV-2 ($p < 0.001$). None of the examined comorbidities had a significant association with COVID-19 diagnosis.

Conclusion

According to the analyses, COVID-Q could be validated since the aspects it evaluated were overall significantly related to SARS-CoV-2 infection. The application of the novel COVID-Q to everyday clinical

practice may help identifying subjects who are likely to be affected by COVID-19 and address them to a nasopharyngeal swab in order to achieve an early diagnosis.

What is already known about this topic?

COVID-19 symptoms are widely known. Lots of studies have been published regarding self-administered questionnaires in order to characterize and know as much as possible regarding this disease. By the way, no specific questionnaires have been validated, yet, and there is no consensus regarding this topic.

What does this article add?

This paper shows the COVID-Q, a novel symptom questionnaire specific for COVID-19 patients. The aim is to provide a comprehensive evaluation that may be helpful to clinicians in order to suspect SARS-CoV-2 infection or not.

Keywords: questionnaire; COVID-19; symptom assessment; SARS-CoV-2; subjective evaluation.

INTRODUCTION

COVID-19 is an emerging infectious disease due to SARS-CoV-2, whose worldwide outbreak led, during late 2019 and during 2020, to the current pandemic condition¹⁻³.

The clinical presentation of COVID-19 ranges from the complete absence of symptoms to a severe acute respiratory distress syndrome, potentially leading to death^{1,4,5}. In symptomatic patients, the most common presentation includes fever, myalgia and fatigue, associated with upper and lower respiratory tract involvement, resulting in nasal congestion, sore throat, anosmia, ageusia, dry cough, and dyspnea⁴. In addition to these respiratory symptoms, gastrointestinal⁶, cardiovascular⁷ and neurological⁸⁻⁹ manifestations have also been described.

Among the upper airways symptoms, the new onset of an altered sense of smell or taste has been regarded with particular interest as clinical hallmark of early SARS-CoV-2 infection, potentially able to discriminate subjects with COVID-19 from those with other types of acute respiratory tract infections¹⁰⁻¹⁷.

Given the critical importance of a comprehensive assessment of the severity of COVID-19 clinical presentation, risk factors and comorbidities useful to characterize patients who will develop a severe COVID-19 have been discussed.¹⁸ On the other hand, currently there is a lack of validated and standardized means to evaluate symptoms.

Clinical evaluation models based on patient-reported outcomes (PROs) have been regarded with increasing interest, in particular in the field of pneumology, to evaluate conditions such as asthma, pneumonia, and chronic obstructive pulmonary disease (COPD)¹⁹⁻²².

To date, only a PRO questionnaire, with adequate content validity and psychometric properties, has been developed and validated to measure the severity, functional impact and response to therapy of acute respiratory tract infections in adults²³. Such reporting tool, the Acute Respiratory Tract Infections Questionnaire (ARTIQ)²³, resulted to be effective in discriminating between participants with and without acute respiratory tract infections.

To the best of our knowledge, no standardized tool has been proposed yet for a uniform and comprehensive evaluation and reporting of symptoms in COVID-19 patients.

The present study aimed to develop and validate the COVID-Q, a novel questionnaire specific for COVID-19 patients in order to provide a comprehensive and standard clinical evaluation. A secondary goal of the present

study was to evaluate the performance of the questionnaire in identifying subjects at higher risk of being tested positive for COVID-19.

MATERIALS AND METHODS

This study was approved by the Ethical Committees of the Academic Hospital of Treviso (Italy) and the Hospital of Belluno (Italy), and it was conducted in accordance with the principles of the Helsinki Declaration. Data were examined in compliance with Italian privacy and sensitive data laws. All participants gave their consent to be enrolled in this investigation and to the treatment of their personal data for scientific purposes. The study was conducted in the Hospital of Treviso.

Questionnaire construction

All patients both in the case and the control group answered the COVID-Q. The COVID-Q included items focused on general, ENT, respiratory and gastrointestinal symptoms, as well as psychological symptoms, drug assumption and changes in daily activity. Such items in part belonged to the ARTIQ²³ and the SNOT-22²⁴ questionnaires, and the remaining were developed *ex novo* by us to better characterize symptoms and clinical conditions specific of COVID-19, basing on the data available from the Atlanta CDC guidelines²⁵. All items based on the ARTIQ were scored on a three-point scale (0 = none, 1 = a little, 2 = a lot) in line with the ARTIQ itself, while the questions focused on anosmia and ageusia, were scored on a scale from 0 (none) to 5 (complete loss of smell or taste), being derived from the SNOT-22. An additional section of the questionnaire focused on demographic data, including age, sex, smoking status (0 = never, 1 = former smoker, 2 = current smoker) and comorbidities (0 = absent, 1 = present).

All subjects in both groups answered the questionnaire via a telephone interview, which took place within 19 days from the first diagnosis in cases and within 43 days from the last negative test in controls. **Participants**

Consecutive subjects who underwent nasopharyngeal and throat swabs for SARS-CoV-2 detection were included in the study. People were tested in the hospital of Treviso between March 4th and 24th, 2020.

Tests of SARS-CoV-2 RNA by polymerase chain reaction on nasopharyngeal and throat swabs were performed according to the World Health Organization recommendation²⁵. Reasons for being tested were symptoms, close contact with affected individuals or working in public health services. Non-hospitalized adults (≥ 18 years) were considered eligible for the study. The whole sample was split into two arms: “cases” who tested positive for SARS-CoV-2 infection and “controls” who tested negative.

The inclusion criteria were the consent to take part to the study and positive test for COVID-19 (for inclusion in the cases group). The main exclusion criteria were clinical conditions leading to the inability to answer the COVID-Q (severe respiratory syndrome, hospitalization in the Intensive Care Unit, cognitive impairment). All subjects in both groups answered the questionnaire via a telephone interview.

Statistical Analyses

The statistical analyses were performed only on 41 of the items from the ARTIQ and SNOT-22, since the remaining regarded self-administrated drugs and social activities, which were not taken into account due to the peculiar circumstances of data collection. Indeed, participants were tested during the first lockdown phase in Italy, when social and work activities had been heavily restricted and people were advised not to take any medication without consulting with their physician.

First, Parallel Analysis and Principal Component Analysis (PCA) were used to identify clusters of items measuring the same dimension. Given the ordinal nature of the data in question, the analyses were performed on the polychoric correlation matrix.

IRT models for polytomous items were fit on each of the components identified by the PCA. We first compared the fit of a freely estimated Generalized Partial Credit Model (GPCM)²⁷ to a common-slope GPCM, in which the slopes (i.e., discrimination parameters) of all items are constrained to be equal. The IRT-based analyses evaluated the functioning of item categories, the presence of clusters of local dependence among items, item fit within the model and model fit to the data. Items were selected through an iterative procedure. Items for which the three response categories did not function as intended were recoded into two categories: 0 = absence of symptom, 1 = presence of symptom. Local dependence was measured by the Q3 index²⁸: following Marais's²⁹ indications, residual correlations between items were compared to the average residual correlation within the dataset, and the greatest deviances were taken to suggest possible item redundancy or underlying dimensions. Item fit was based on the scaled version of the chi square statistic χ^2_{*30} : p-values smaller than .05 were considered as indicating significant misfit and led to the exclusion of the item. Finally, omnibus fit was

evaluated by the M2* limited-information test³¹ and was considered acceptable when p values were above the .05 threshold. The items retained in the IRT models were then used to produce scale and total scores, computed as raw sums, and represented the final validated version of the questionnaire.

We estimated the extent to which self-reported symptoms could help identify COVID-19 by means of two different logistic regression models. In both models, the diagnosis (1 = positive, 0 = negative) was the dependent variable, and age, sex, smoking status, and presence of comorbidities were used as covariates. The predictors of interest were the total questionnaire score in the first model, and individual scale scores in the second model. By design, affected individuals were over-represented in our sample compared to the population: this was taken into account by applying the appropriate correction for rare events³². Further examinations were performed according to the obtained results to refine items by using the DETECT method in package sirt^{33,34}. Finally, we identified questionnaire scores associated with specific risk thresholds for COVID-19.

RESULTS

Sample characteristics

A sample of 460 participants (176 males and 286 females; mean age = 51 years, range 20-99 years) was recruited. Among participants there were 230 cases (113 males, and 117 females; mean age 55 years, range 20-99 years) and 230 controls (61 males, and 169 females; mean age 46 years, range 21-89). The telephone interview took place within 19 days from the first diagnosis in cases and within 43 days from the last negative test in controls. The latter time frame resulted to be reliable because, being the controls negative, there was no evolution in their clinical condition, which was comparable with the moment of the first swab.

The prevalence of smoke and comorbidities in cases and controls is shown in **Supplementary Table S1**. An independent sample t-test revealed that the average number of comorbidities (as considered as present or not present) per person was significantly higher among cases than controls ($t(11) = 2.263$, $p = .045$). In particular, Fisher's exact tests showed that hypertension, diabetes, and cardiovascular diseases were significantly more frequent among cases, compared with controls ($p < .001$, $p = .001$ and $p = .030$, respectively). No significant differences in the distribution of the remaining comorbidities or in the proportion of smokers emerged between the two groups. Liver diseases were excluded from all analyses, having been observed in only one case.

Principal component analysis

Parallel analysis led to the extraction of five components, which corresponded to as many clinical presentation patterns: asthenia (including fatigue, sleep disorders and depression mood), influenza-like symptoms (including fever, stomachache and headache), ear and nose symptoms, breathing issues, and throat symptoms. Two symptom classes (chest pain and anosmia/ageusia), did not load clearly on any of the principal components. Chest pain was therefore excluded from further analyses, while the symptom of anosmia/ageusia, given its relevance to COVID-19 symptomatology, was retained as a single item, separated from the other five symptom classes. The five detected components shared weak to moderate correlations (r s between 0.15 and 0.52) and together accounted for 71.05 % of the observed variance.

Item Response Theory models

The unconstrained GPCM for Asthenia, Influenza, Breathing issues, and Throat symptoms showed the best fit to their relative data, according to AIC, BIC and Log Likelihood. Only in the case of ear/nose symptoms, the

fixed-slope (i.e. 1PL) GPCM model showed comparatively better fit to the data and thus was preferred. The comparative statistics of the full and constrained models are presented in **Supplementary Table S2**.

Of the original nine items regarding asthenia, items “vertigo” and “felt dizzy” were removed because of significant misfit to the model. The model estimated on the remaining seven items did not fit the data adequately. The Q3 index showed a high residual correlation – compared to the dataset’s average – between item “been in a bad mood” and “been irritable”, suggesting the presence of a secondary dimension associated with mood alterations. Once these items were removed, leaving a final set of five items, good model-data fit was achieved.

The items regarding influenza-like symptoms did not fit a single GPCM model. Further examination using the DETECT method in package sirt resulted in the presence of a “fever” cluster (including only three items: “feeling feverish”, “sweat” and “chills”), distinct from the remaining 5 items, which assessed more specifically gastro-intestinal issues. We decided to consider these two clusters separately. The Item Response Theory (IRT) model of the “gastro-intestinal” cluster led to dichotomizing item “vomiting” and excluding item “loss of appetite” for misfit, thus resulting in a four-item scale. Due to the very low number of items, we did not fit an IRT model to the fever cluster. Regarding ear and nose symptoms, despite no significant item misfit or problematic category functioning, the initial model did not adequately fit the data. Examination of residual correlations showed high residuals between items “muscle pain” and “joint pain” and between items “painful pressure in the ears” and “painful sinuses”. We therefore excluded items “joint pain” and “painful pressure in the ears” to reduce redundancy. Item “muscle pain” no longer fit the reduced model and was excluded in turn. However, since “muscle or body aches” are indicated as one of the most common symptoms of COVID-19, we decided to retain it as a single item next to the final five items assessing ear and nose symptoms.

Regarding breathing issues, item “difficulty in thinking clearly” and “difficulty in going about your daily business” were removed due to misfit, whereas item “being so unwell you had to stay in bed” was removed due to high residual correlations with item “felt tired”. Item “coughing up mucus” was instead dichotomized to improve the functioning of answer categories. The final model included six items.

All five items associated to throat symptoms fit the model well, and the model offered a good representation of the data.

The final IRT models retained 27 items: 5 assessing asthenia, 3 assessing fever, 4 assessing gastro-intestinal symptoms, 5 assessing nose symptoms, 6 assessing breathing problems and 4 assessing throat symptoms. Items “coughing up mucus” and “vomiting” were recoded as dichotomous (0 = symptom absent. 1 = symptom present).

Items “anosmia/ageusia” and “muscle pain” were not included in any of the models but were kept as single items due to their relevance in diagnosing COVID-19.

Table 1 shows item parameters and item fit index for the final models. **Table 2** shows model fit indices.

Cronbach’s alpha was between 0.69 and 0.92 for all subscales and for the total scale.

Table 3 reports the questionnaire in its final validated form, which contains only the items retained in the IRT models.

Questionnaire scores and risk of COVID-19

Supplementary Table S3 reports descriptive statistics for variables used in the logistic linear regressions. The effects of the total questionnaire score in predicting COVID-19 diagnosis are reported in **Supplementary Table S4**. Results from the model examining individual scale scores (asthenia, gastrointestinal symptoms, fever, ear/nose symptoms, breathing issues, throat symptoms, anosmia/ageusia, and muscle pain) appear in **Supplementary Table S5**.

The total score on the questionnaire was significantly associated with positivity at molecular test for SARS-CoV-2: as expected, individuals displaying a larger number of symptoms or experiencing them more often, as reported in the questionnaire, were significantly more likely to be affected by COVID-19 ($\exp(\beta) = 1.203$, $p < .001$). Age (being older) and sex (being male) also represented risk factors ($\exp(\beta) = 1.038$ and 3.057 , respectively, $p < .001$). Everything else being equal, none of the comorbidities examined had a significant association with COVID-19 diagnosis.

Investigation of the effects of individual scale scores (**Supplementary Table S5**) revealed that breathing difficulties and neurological symptoms of anosmia/ageusia were significantly associated to higher probabilities of COVID-19 diagnosis, when controlling for other variables in the model. Asthenia, gastrointestinal, fever, Ear/Nose and Throat symptoms, instead, did not have a significant impact, everything else being equal. This means that, in the case of e.g. wheezing, shortness of breath, cough (i.e. breathing issues), loss of smell or taste

(i.e. anosmia/ageusia), a greater number or severity of symptoms are associated with an increased chance of testing positive for SARS-CoV-2 swab. The other recorded symptoms, on the other hand, have no impact on the probability of the swab outcome beyond what is already accounted for by the breathing and anosmia/ageusia items. According to this finding, we computed an additional partial score, consisting in the unweighted sum of the breathing symptoms scale and of the anosmia/ageusia item, and used it to estimate scores in association with specific risk thresholds. Results from the model using this partial score as predictor are presented in **Table 4**.

DISCUSSION

. According to the analyses of the present study, six subscales and two individual items, each accounting for a separate dimension of COVID-19 symptomatology, were identified in order to validate the novel questionnaire. Such questionnaire presented some analogy with the ARTIQ²³, although some clinical categories of the latter, such as “ENT” and the “influenza” symptoms, in our novel assessment model turned into more homogeneous components.

Besides providing a reliable tool for assessing and reporting symptom, the COVID-Q also provided a quantitative score which appeared to correlate with the risk of testing positive to SARS-CoV-2 infection. In particular, higher scores on the “breathing issues” subscale (i.e. “dry cough”, “coughing up mucus”, “problems with breathing”, “wheezing”, “shortness of breath”, “felt tired”) and on the anosmia/ageusia item (i.e. “loss of smell or taste”) were significantly associated with a higher chance of a positive SARS-CoV-2 test.

Figure 1 shows the average probability of COVID-19 diagnosis as a function of total score in the breathing symptoms scale and in the anosmia/ageusia item, sex and age group. According to our model, a middle-aged person (i.e. 51 years old, the average age in our sample), without any breathing or neurological symptoms had a 0.63% baseline chance of testing positive to a COVID-19 swab, irrespective of gender, smoking behavior or associated illnesses. This baseline chance increased by an average of 1.99 times for every point scored in either the breathing scale or the anosmia/ageusia item. Thus, a score of 1 on the partial subscale represented a 1.24 % chance of a positive swab, which increased to 16.42% for a score of 5 and 99.74% for a score of 16, the highest possible. These risk thresholds were computed for individuals aged 51 years old and did not take into

account their sex. However, our model predicted that sex and age would further affect the baseline probability. Namely, the chance of testing positive to a swab increased or decreased by 1.04 times for every year of age over or below 51, respectively. Similarly, males were 2.12 times more likely to be affected by COVID-19, compared to the baseline probabilities presented above, while females were 1.58 times less likely. These results are in line with those reported in the literature, since male gender, for instance, has been associated with worse course of COVID-19, up to higher risk of hospitalization in Intensive Care Unit and even risk of death.³⁶ Moreover, regarding age, a similar conclusion has been previously obtained, being immunosenescence and inflammaging factors which may determine a less favourable course in older adults.³⁷

It must be borne in mind that great care must be exercised to ensure an adequate description of the patient's condition and symptoms, especially in a field, like the emerging COVID-19, where no gold standard exists. The questionnaire presented in this study has been used in previous published articles¹³⁻¹⁶ and allowed to obtain reliable results regarding patients' symptoms. In those papers, indeed, a proper assessment could be obtained to follow-up a group of patients who had previously tested positive for SARS-CoV-2. Many other available studies on COVID-19 were focused on patient-reported symptoms, also collected via non-conventional means of clinical assessment, including tele-medicine³⁸ and the use of social networks³⁹. However, to date, a comprehensive and specific tool to collect and scale symptoms from COVID-19 patients has not been described yet.

The COVID-Q may potentially fill this gap, providing a validated and standardized assessment tool to support clinical data collection, also in the absence of a direct contact with the patient, such as in the tele-medicine field.

Although this study also provided score thresholds to define patients at higher risk of testing positive to the SARS-CoV-2 infection, the COVID-Q should not be used as a diagnostic tool. Rather, its application to those who did not receive a microbiological test may help to discriminate COVID-related symptoms, potentially aiming to identify subjects who should be tested. To the best of our knowledge, no validated questionnaires have been developed so far. Usefulness of the questionnaire shown in this paper relies in the high possibility to discriminate the probability of a patient to test positive to SARS-CoV-2. Given the case of a subject who can't undergo nasopharyngeal swab (e.g., logistical reasons, home bedridden patients), administration of COVID-Q may be a possible and reliable solution for General Practitioners to suggest self-isolation.

The main strengths of this study reside in the controlled design, which allowed to apply the questionnaire to both COVID-19 patients and healthy subjects, and in the accurate definition of its items. The item response theory models led to the development of a panel of questions which allowed to obtain unbiased descriptions of clinical condition, questions focused to meet COVID-19-related symptoms, and a reliable scaling system. On the other hand, it has to be considered that the COVID-Q has been developed and applied in Italian language. Any other language version should be tested to verify whether it retained its measurement properties also in different linguistic and cultural or ethnic backgrounds. Another limitation of the present study is the number of patients. Even if it's statistically significant, a wider application of the COVID-Q to a larger number of patients may allow to obtain even more precise results and, possibly, to adjust the model we suggested with the present study. It must be observed, also, that being SARS-CoV-2 an RNA virus, it can mutate into diverse variants which may give diverse presentations of COVID-19. Therefore, possible score differences may be encountered by using the questionnaire. At last, the present study has been conducted in a single geographical area, so its application in different areas of other Countries may be useful in order to observe any possible score differences.

In conclusion, in this study we developed and proposed a novel questionnaire to assess and scale the COVID-19-associated symptoms. Its application to COVID-19 patients could improve the reliability in collecting clinical data and assessing the severity of symptoms, while its extension to subjects who have not received a COVID-19 diagnosis yet might help to identify those at higher risk of being affected and, therefore, needing a microbiological test.

REFERENCES

1. Rothan, H.A. & Byrareddy, S.N. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. *J Autoimmun* **109**,102433 (2020).
2. Lu, R. et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565-574 (2020).
3. Wu, Y.C., Chen, C.S. & Chan, Y.J. The outbreak of COVID-19: An overview. *J Chinese Med Assoc* **83**, 217-220 (2020).
4. Lake, M.A. What we know so far: COVID-19 current clinical knowledge and research. *Clin Med* **20**, 124-127 (2020).
5. Huang, C., et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497-506 (2020).
6. Tian, Y., Rong, L., Nian, W. & He, Y. Review article: gastrointestinal features in COVID-19 and the possibility of faecal transmission. *Aliment Pharmacol Ther* **51**, 843 (2020).
7. Driggin, E., et al. Cardiovascular Considerations for Patients, Health Care Workers, and Health Systems During the COVID-19 Pandemic. *J Am Coll Cardiol* **75**, 2352-2371 (2020).
8. Asadi-Pooya, A.A. & Simani, L. Central nervous system manifestations of COVID-19: A systematic review. *J Neurol Sci* **413**, 116832 (2020).
9. Krajewska, J., Krajewski, W., Zub, K. & Zatoński, T. COVID-19 in otolaryngologist practice: a review of current knowledge. *Eur Arch Otorhinolaryngol* **277**, 1885-1897 (2020).
10. Klopfenstein, T., et al. Features of anosmia in COVID-19. *Med Mal Infect* **50**, 436-439 (2020).
11. Yan, C.H., et al. Association of chemosensory dysfunction and COVID-19 in patients presenting with influenza-like symptoms. *Int Forum Allergy Rhinol*, **10**:806-813 (2020).
12. Fabbris, C., et al. Is oro/nasopharyngeal swab for SARS-CoV-2 detection a safe procedure? Complications observed among a case series of 4876 consecutive swabs. *Am J Otolaryngol* **42**,102758 (2021).
13. Boscolo-Rizzo, P., et al. Predominance of an altered sense of smell or taste among long-lasting symptoms in patients with mildly symptomatic COVID-19. *Rhinology* **58**, 524-525 (2020).
14. Boscolo-Rizzo, P., et al. Evolution of Altered Sense of Smell or Taste in Patients With Mildly

- Symptomatic COVID-19. *JAMA Otolaryngol Head Neck Surg* **146**, 729–32 (2020).
15. Spinato, G., et al. Alterations in Smell or Taste in Mildly Symptomatic Outpatients With SARS-CoV-2 Infection. *JAMA* **323**, 2089-2090 (2020).
16. Boscolo-Rizzo, P., et al. New onset of loss of smell or taste in household contacts of home-isolated SARS-CoV-2-positive subjects. *Eur Arch Otorhinolaryngol*. **277**, 2637-2640 (2020).
17. Borsetto, D., et al. Self-reported alteration of sense of smell or taste in patients with COVID-19: a systematic review and meta-analysis on 3563 patients. *Rhinology* **58**, 430-436 (2020).
18. Del Sole, F., Farcomeni, A., Loffredo, L., Carnevale, R., Menichelli, D., Vicario, T., Pignatelli, P., & Pastori, D. Features of severe COVID-19: A systematic review and meta-analysis. *European journal of clinical investigation* **50**, e13378 (2020).
19. Barrett, B., et al. The Wisconsin Upper Respiratory Symptom Survey (WURSS): a new research instrument for assessing the common cold. *J Fam Pract* **51**, 265 (2002).
20. Juniper, E.F., et al. Development and validation of a questionnaire to measure asthma control. *Eur Respir J* **14**, 902–907 (1999).
21. Lamping, D.L., et al. The community-acquired pneumonia symptom questionnaire: a new, patient-based outcome measure to evaluate symptoms in patients with community-acquired pneumonia. *Chest* **122**:920–929 (2002).
22. Meguro, M., Barley, E.A., Spencer, S. & Jones, P.W. Development and validation of an improved, COPD-specific version of the St. George Respiratory Questionnaire. *Chest* **132**,456–63 (2007).
23. Aabenhus, R., Thorsen, H., Siersma, V., Brodersen, J. The development and validation of a multidimensional sum-scaling questionnaire to measure patient-reported outcomes in acute respiratory tract infections in primary care: The acute respiratory tract infection questionnaire. *Value Heal* **16**, 987-992 (2013).
24. Hopkins, C., et al. Psychometric validity of the 22-item Sinonasal Outcome Test. *Clin Otolaryngol* **34**, 447-454 (2009).
25. <https://www.cdc.gov/coronavirus/2019-ncov/symptoms-testing/symptoms.html>.
26. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance-publications?publicationtypes=01bc799c-b461-4a52-8c7d-294c84cd7b2d>

27. Muraki, E. A generalized partial credit model: Application of an EM algorithm. ETS Research Report Series **1**, i-30 (1992).
28. Yen, W.M. Effects of Local Item Dependence on the Fit and Equating Performance of the Three-Parameter Logistic Model. *Applied Psychological Measurement* **8**, 125-145 (1984).
29. Marais, I. *Local Dependence* in Christensen, K.B., Kreiner, S. & Mesbah, M. (Eds), *Rasch models in health* (JohnWiley & Sons, Hoboken, NJ, 2013).
30. Stone, C.A. Monte Carlo based null distribution for an alternative goodness-of-fit test statistic in IRT models. *J Educ Meas* **37**, 58-75 (2000).
31. Cai, L. & Hansen, M. Limited-information goodness-of-fit testing of hierarchical item factor models. *Br J Math Stat Psychol* **66**, 245-276 (2013).
32. King, G. & Zeng, L. Logistic regression in rare events data. *J Stat Softw* **8**, 137-163 (2003).
33. Zhang, J. & Stout, W. Dimensionality-Based Partition of a Test, Where the Number of Item Clusters in This Partition Corresponds To the Number of Substantively. *Psychometrika* **64**, 213-249 (1999).
34. Robitzsch, A. sirt: Supplementary Item Response Theory Models. R package version 3.10-70. <https://CRAN.R-project.org/package=sirt> (2020)
35. Cai, L. & Monroe, S. *A New Statistic for Evaluating Item Response Theory Models for Ordinal Data* in *CRESST Report 839* (University of California, National Center for Research on Evaluation, Standards, and Student Testing –CRESST-, Los Angeles, California, 2014).
36. Peckham, H., de Gruijter, N. M., Raine, C., Radziszewska, A., Ciurtin, C., Wedderburn, L. R., Rosser, E. C., Webb, K., & Deakin, C. T. Male sex identified by global COVID-19 meta-analysis as a risk factor for death and ICU admission. *Nature communications* **11**, 6317 (2020).
37. Pietrobon, A. J., Teixeira, F., & Sato, M. N. Immunosenescence and Inflammaging: Risk Factors of Severe COVID-19 in Older People. *Frontiers in immunology*, **11**, 579220 (2020).
38. Mann, D.M., et al. COVID-19 transforms health care through telemedicine: Evidence from the field. *J Am Med Inform Assoc* **27**, 1132-1135 (2020).
39. Sarker, A., et al. Self-reported COVID-19 symptoms on Twitter: an analysis and a research resource. *J Am Med Inform Assoc* **27**, 1310-1315 (2020).

FIGURE LEGEND

Figure 1. Probability of testing positive to COVID-19 swab (“p” on y axis) as a function of total score on the COVID-Q breathing scale and anosmia/ageusia (AG/AN) item (x axis), age group and sex. Note: probability was computed for the mean age in each age group.