Predictive value of olfactory and taste symptoms in the diagnosis of COVID-19: A systematic review and meta-analysis

Do Hyun Kim, M.D., Ph.D.,[1] Sung Won Kim, M.D., Ph.D.[1] Gulnaz Stybayeva, M.D., Ph.D.,[2] So Yun Lim, M.D.,[1] Se Hwan Hwang, M.D., Ph.D.,[3*]

[1]*Department of Otolaryngology-Head and Neck Surgery, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea.*

[2]*Department of Physiology and Biomedical Engineering, Mayo Clinic, Rochester, MN, USA*

[3]*Department of Otolaryngology-Head and Neck Surgery, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea.*

Running title: Symptoms to diagnose COVID-19

ORCID: Do Hyun Kim, 0000-0002-9248-5572; Sung Won Kim, 0000-0002-8981-2536; Gulnaz Stybayeva, 0000-0002-1453-245X; So Yun Lim, 0000-0001-8017-7890; Se Hwan Hwang, 0000-0002-2838-7820

Conflict of interest: None

Sponsorships: None

Funding source: This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A1B07045421), the Bio & Medical Technology Development Program of the NRF funded by the Ministry of Science & ICT(2018M3A9E8020856, 2019M3A9H2032424, 2019M3E5D5064110). The sponsors had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions: Conceptualization: DHK, SHH, Data curation: SWK, SG, SYL, Formal analysis: SWK, SG, SYL, Funding acquisition: DHK, SHH, SWK, Methodology: SWK, SG, SYL, Project administration: DHK, SHH. Visualization: DHK, SHH. Writing - original draft: DHK, SHH. Writing - review & editing: DHK, SHH, SWK, SG, SYL.

***Corresponding author:** Se Hwan Hwang, M.D., Ph.D.

Department of Otolaryngology-Head and Neck Surgery, Bucheon St. Mary's Hospital, College of Medicine, The Catholic University of Korea.

Address: 327 Sosa-ro, Bucheon-si, Gyeonggi-do, 14647, Korea.

Phone: +82 32 340 7044; Fax: +82 32 340 2674; E-mail address: yellobird@catholic.ac.kr

Abstract

Objectives: This study evaluated the diagnostic value of the various symptoms of COVID-19 in the screening of this disease.

Methods: Two authors (working independently) comprehensively reviewed six databases (PubMed, Cochrane database, Embase, Web of Science, SCOPUS, and Google Scholar) from their dates of inception until November 2020. Patient-reported symptoms, including otolaryngologic and general symptoms, were evaluated for their predictive values in adults who underwent testing for COVID-19. True-positive, true-negative, false-positive, and false-negative data were extracted from each study. The methodological quality of included studies was evaluated using the Quality Assessment of Diagnostic Accuracy Studies tool (ver. 2).

Results: Twenty-eight prospective and retrospective studies were included in the meta-analysis. The diagnostic odds ratio (DOR) of a change in olfaction and/or taste was 10.20 (95% confidence interval [CI], 8.43; 12.34). The area under the summary receiver operating characteristic curve was 0.8. Olfactory and/or taste changes had a low sensitivity (0.57, 95%CI: 0.47; 0.66) but moderate negative (0.78, 95%CI: 0.69; 0.85] and positive (0.78, 95%CI: 0.66; 0.87) predictive values and a high specificity (0.91, (95%CI: 0.83; 0.96). Olfactory and/or taste changes had a higher diagnostic value than the other otolaryngologic symptoms, a higher DOR and specificity, and a similar or higher diagnostic value than the other general symptoms.

Conclusions: Among otolaryngologic symptoms, olfactory and/or taste dysfunction was the most highly associated with COVID-19 and its general symptoms and should be considered when screening for the disease.

Keywords: Coronavirus Infections; Olfaction Disorders; Ageusia; Dysgeusia

HIGHLIGHTS

- It is important to timely predictive symptoms of COVID-19 for quarantining of patients.
- Olfactory and/or taste dysfunction was the most highly associated with COVID-19.
- Validated olfactory and/or taste tools have higher diagnostic value for COVID-19.

Introduction

Since its initial outbreak in 2019, the acute respiratory illness caused by the SARS-CoV-2 virus (COVID-19) continues to spread at an exponential rate. The causative agent, the newly discovered coronavirus, is most frequently transmitted between people through respiratory droplets and aerosols [1]. Influenza-like symptoms or mild pneumonia develop in > 80% of COVID-19 patients, such that most patients do not need to be hospitalized [2]. However, significant viral transmission has been traced to these mildly symptomatic and non-hospitalized patients [3].

In the absence of a specific treatment and with vaccine trials still underway [4], the rapid and reliable diagnosis of COVID-19 infection followed by the strict isolation of patients is the most effective means of controlling disease spread [5]. Currently, the diagnosis of COVID-19 is mostly made by RT-PCR testing of respiratory samples, with further discriminative features of the disease often apparent on chest CT scans [6,7]. However, RT-PCR tests are not always readily available, especially in some countries or regions, and the delayed reporting of test results due to the large number of samples in certain institutions may lead to a delay in the proper quarantining of patients. Since the outbreak of the pandemic, the clinical symptoms of COVID-19-positive patients have been described in many reports [8]. Given the limited clinical resources, it is important to identify the most predictive symptoms of COVID-19 infection to ensure the timely quarantining of patients and thereby disease spread [9]. Therefore, we conducted a meta-analysis comparing the diagnostic value of olfactory and/or taste changes as well as other otolaryngologic symptoms and general symptoms with the current reference test (RT-PCR). Also, considering the inclusion of various and heterogeneous studies, the diagnostic accuracy of COVID-19 was sub-analyzed

according to the validation of olfactory and/or taste disorder (OTD) questionnaires or tools as well as demographic factors and severity of disease.

Materials and Methods

Ethical statements

This review study did not treat human participants. Therefore, our Institutional Review Board waived the need for informed consent for a systematic review and meta-analysis.

Literature search

Clinical studies were retrieved from PubMed, the Cochrane Central Register of Controlled Trials, Embase, Web of Science, SCOPUS, and Google Scholar. The search period was from the date of database inception until November 2020. The search terms were: “coronavirus disease 2019”, “severe acute respiratory syndrome coronavirus 2”, “coronavirus”, “COVID-19”, “anosmia”, “ageusia”, “dysgeusia”, “smell”, “taste”, “smell disorders”, “taste disorders”, “PCR”, “diagnostic accuracy”, “signs”, “symptoms”, “cough”, “diarrhea”, “dyspnea”, “fatigue”, “fever”, “headache”, “myalgia”, “fatigue”, and “fever”. Only studies written in English were reviewed. When we performed five database searches, the keywords were used by the combinations (“or”) of all possible keywords ([all fields] and the language limitation such as English (“and”). For very brief and partial example, the following combination of search details was used in MEDLINE: ("COVID-19"[Mesh]”, or “coronavirus disease 2019”[All Fields], or “severe acute respiratory syndrome coronavirus 2” [All Fields]) AND “diagnosis”[All Fields] AND ("Signs” [All Fields] and “Symptoms[All Fields] OR (“anosmia” [Mesh] OR "Smell"[Mesh] OR "Olfaction Disorders"[Mesh] OR “Ageusia"[Mesh] OR

"Dysgeusia"[Mesh] "Taste"[Mesh] OR "Taste Disorders"[Mesh] OR "Taste and Smell Impairment" [All Fields]) OR ("Cough"[Mesh] OR "Cough" [All Fields] OR "Coughs" [All Fields]) OR ("Diarrhea"[Mesh] OR "Diarrhea" [All Fields] OR "Diarrheas" [All Fields]) OR "Fatigue"[Mesh] OR "Fatigue" [All Fields] OR "Lassitude" [All Fields]) OR ("Fever"[Mesh] OR "Fever" [All Fields] OR "Fevers" [All Fields] OR "Pyrexia" [All Fields] OR "Pyrexias" [All Fields]) OR "Headache"[Mesh] OR "Myalgia"[Mesh] AND English[All Fields]). We used similar search words for the other databases. The reference lists of each publication were examined to ensure that no relevant studies had been omitted. All abstracts and titles of candidate studies were assessed by two independent reviewers. Studies that did not address smell and taste disorders in the context of COVID-19 were excluded. Detailed search terms and queries was described in Supplementary Table 1.

Selection criteria

The inclusion criteria were: 1) English language, 2) prospective or retrospective study protocol, 3) comparison of the prevalence of various symptoms, including smell or taste disorders, in patients or controls tested by PCR via pharyngeal swab, and 4) eligibility in sensitivity and specificity analyses. The exclusion criteria were: 1) case report format, 2) review article format, and 3) lack of diagnostic power regarding smell or taste disorders. The search strategy is summarized in Figure 1.

Data extraction and risk of bias assessment

We compared the results of the various symptoms with the results of the PCR from respiratory secretions. Then, we extracted TP (true positive), FP (false positive), TN (true

negative), and FN (false negative) values to calculate diagnostic accuracy, defined as the diagnostic odds ratio (DOR), sensitivity, specificity, negative predictive value, and positive predictive value. The calculation was as follows: DOR, (TP/FP)/(FN/TN); sensitivity, TP /(TP+FN); specificity, TN/(TN+FP); negative predictive value, TN/(TN+FN); positive predictive value, TP/(TP+FP). Summary receiver operating characteristic (SROC) curve and the area under the curve (AUCs) were also analyzed together[1-6,9-42].

DORs were calculated with 95% confidence intervals (CIs), using random-effects models that considered both within- and between-study variation. DOR values ranged from 0 to infinity, with higher values indicative of a better diagnostic performance. A value of 1 indicated that the presence or absence of disease could not be inferred. The logarithm of each DOR was calculated to obtain an approximately normal distribution [43]. The SROC approach is the method of choice for the meta-analysis of studies reporting both sensitivity and specificity. As the discriminatory power of a test increases, the SROC curve shifts toward the top left-hand corner of the ROC space (i.e., toward the point where both sensitivity and specificity equal 1 [100%]). The AUC can range from 0 to 1, with higher values indicative of a better performance. To be useful, a diagnostic tool must exhibit good reliability; thus, our analysis focused on the reliability of symptoms. As the data were examined by clinicians, the most important type of reliability was interrater agreement, assessed by comparing interpretations of the results between two or more independent assessors. From all studies, data were collected regarding the number of patients, the true-positive, true-negative, false-positive, and false-negative values, which were used to calculate the AUCs and the DORs. Study quality was analyzed using the Quality Assessment of Diagnostic Accuracy Studies tool (ver. 2; QUADAS-2).

Statistical analysis and outcome measurements

The meta-analysis was conducted using the R statistical software (R Foundation for Statistical Computing, Vienna, Austria). The R package MADA was used to perform the pooling of diagnostic outcomes and generate SROC curves. Pooled sensitivity, specificity, negative predictive value (NPV), positive predictive value (PPV), DOR outcomes were generated, with 95% CI. Heterogeneity, referring to the variation in study outcomes between studies, was then analyzed using I^2. The measure ranged from 0 (no heterogeneity) to 100 (maximum heterogeneity). Those outcomes that did not present a significant level of heterogeneity ($I^2 < 50$) were analyzed with the fixed-effects model. It is assumed that all studies come from a common population. By contrast, when significant heterogeneity among outcomes was found (defined as $I^2 > 50$), the random-effects model was used. This model assumes that the true effects in individual studies may be different from one another, and that these are normally distributed. Forest plots were drawn for the sensitivity, specificity, and negative predictive value and for the SROC curves.

Results

Thirty-eight studies comprising 120,256 participants were included in this meta-analysis. The study characteristics were described in Supplementary Table 2, and bias assessment of the studies were shown in Supplementary Table 3.

Diagnostic accuracy of OTD and only olfactory disorder (OD)

Eleven prospective and retrospective studies addressing OTD were included. The DOR of

OTD was 10.20 (8.43; 12.34, I^2 = 64.0%) (Figure 2). The area under the SROC curve was 0.80 (Figure 3). OTD had a low sensitivity (0.57, 95%CI: 0.47; 0.66, I^2 = 97.5%) but moderate negative (0.78, 95%CI: 0.69; 0.85, I^2 = 98.7%) and positive (0.78, 95%CI: 0.66; 0.87, I^2 = 98.7%) predictive values and a high specificity (0.91, 95%CI: 0.83; 0.96, I^2 = 99.4%) (supplementary Figure 1).

Seventeen prospective and retrospective studies addressed OD. The DOR of OD was 10.37 (95%CI: 6.31; 17.05, I^2 = 83.9%) (Figure 2), and the area under the SROC curve 0.80 (Figure 3). An olfactory test alone yielded similar results to OTD with respect to its diagnostic accuracy, with a low sensitivity (0.50, 9%%CI: 0.34; 0.66, I^2 = 97.1%), moderate negative (0.77, 95%CI 0.64; 0.87, I^2 = 98.8%), and positive (0.78, 95%CI: 0.66; 0.87, I^2 = 93.8%) predictive values, and a high specificity (0.93, 95%CI: 0.86; 0.97, I^2 = 97.2%) (supplementary Figure 2). Compared with OTD, OD had a lower sensitivity (0.50 vs 0.55, p<0.001) but a higher specificity (0.93 vs 0.91, p =0.0003) and DOR (10.20 vs 10.37, p <0.001). By contrast, there were no significant difference of NPV (0.77 vs 0.78, p =0.065) and PPV (0.78 vs 0.78, p = 0.82) in both groups.

Given the statistical heterogeneity in the accuracy of the diagnosis, both the heterogeneity and the diversity of the enrolled studies had to be taken into account to ensure that there were no significant biases. Thus, a subgroup analysis was performed to analyze the effects of the different measurements of olfactory or taste dysfunction (validated instruments vs. nonvalidated survey), severity of COVID-19 symptoms (mild to moderate vs severe), and ethnicity (Asian vs Caucasian) on the diagnostic efficacy.

For the OTD data, the validated instruments subgroup comprised only one study, such that a

subgroup analysis was not possible. For the OD data, the validated instruments subgroup consisted of three studies, which were then subjected to a subgroup analysis. The validated instruments subgroup tended to be less specific (0.92 vs. 0.93) but the sensitivity (0.79 vs. 0.44), NPV (0.83 vs. 0.76), PPV (0.85 vs. 0.75), and DOR (41.30 vs. 9.02) were higher than in the non-validated instruments subgroup. In subgroups analysis regarding severity of COVID-19 symptoms, the severe subgroup tended to be less sensitive (0.37 vs 0.61; 0.41 vs 0.59) and less predictable (0.52 vs 0.82; 0.77 vs 0.78) but more specific (0.98 vs 0.89; 0.97 vs 0.87) and more predictable (0.94 vs 0.73; 0.85 vs 0.74) than mild to moderate subgroup in the OTD and OD, respectively. For the subgroup analysis regarding the ethnicity was not possible, because only one study covered the Asian group.

Diagnostic accuracy of other otorhinologic symptoms and general symptoms

Other otorhinolaryngologic symptoms, such as nasal symptoms and sore throat, were of low diagnostic accuracy (sensitivity: 20%, specificity: 74~80%, NPV: 62–75%, PPV: 22~30%, AUC:0.46~54). There were no significant associations between these symptoms and the prevalence of COVID-19. However, sore throat (DOR, 0.66, 95%CI: 0.38; 1.15) tended to be negatively related to +COVID-19.

Among the generalized symptoms (cough, diarrhea, dyspnea, fatigue, fever, headache, and myalgia), diarrhea, fatigue, fever, and myalgia were significantly positively correlated with COVID-19 positivity. Diarrhea and dyspnea were of low sensitivity (0.10–0.20) and PPV (0.20–0.30), and of moderate specificity and NPV (0.70–0.80). Fatigue, fever, and myalgia were of moderate specificity (0.5–0.8) and NPV (0.7–0.8) and of low sensitivity (0.4–0.6) and PPV (0.2 –0.3). Thus, other symptoms were diagnostically less powerful than OTD

(Table 2).

Discussion

The early and accurate diagnosis of SARS-CoV-2 infection is key to halting the COVID-19 pandemic, given the high propagation rate of the virus, the rapid spread of disease worldwide, and the adverse, often fatal consequences of infection [1,6]. The autumn-winter season in the northern hemisphere is generally marked by the circulation of influenza and other respiratory viruses that initially may be difficult to distinguish from COVID-19 [17]. While RT-PCR and thoracic CT scan are definitive diagnostic tools, their accessibility may be limited due to a shortage of medical resources or inefficient policy-making decisions. Thus, control the spread of COVID-19 in the community requires that the distinctive clinical features of the disease be readily recognized such that those patients can then be appropriately managed [18].

Currently, the COVID-19 symptoms recognized by the World Health Organization include coughing, fever, fatigue, and difficulty breathing [1]. The U.S. Centers for Disease Control and Prevention initially listed list three major symptoms: fever, cough, and shortness of breath, but as the epidemic progressed added chills, myalgias, headache, sore throat, and the loss of taste and/or smell [14]. However, the clinical manifestation of patients with COVID-19 are often non-specific, resembling those of other influenza-like illnesses and thus complicating a clinical diagnosis of COVID-19. As data regarding the diagnostic power of highly specific symptoms in predicting COVID-19 positivity are limited [1], we quantified the specificity, sensitivity, PPV, and NPV of symptoms reported by the WHO and the health authorities of other countries in a pooled sample of patients who underwent SARS-CoV-2

testing, including those with positive and negative test results. This information is important for both the general public and health care professionals, as it enables faster and more effective isolation procedures and treatment [10].

In our study, OTD had a pooled sensitivity of 0.57, a pooled specificity of 0.91, a pooled NPV of 0.78, a pooled PPV of 0.78, and an AUC of 0.80. The area under the SROC (0.70–0.80) indicated moderate diagnostic accuracy [44]. The sensitivity of OTD in detecting COVID-19 positivity was 56%, which is not high enough for diagnostic purposes. However, the specificity of OTD in estimating COVID-19 negativity was 90%, which is high enough to exclude false-positive COVID-19 diagnoses. In a direct comparison of OTD with OD, OTD was less specific (0.93 vs. 0.9107, $p < 0.001$) but more sensitive (0.50 vs. 0.55, $p < 0.001$). These results showed that, for patients with apparent COVID-19, there is no clinical difference between OTD and OD.

Based on the negative and positive predictive values determined in this study (70–80%), false-negatives and false-positives would need to be considered in the use of OTD and OD to detect COVID-19. With an NPV of ~80%, a negative test would be a false-negative in 20% of the patients and COVID-19 would therefore go undetected. Conversely, a PPV of ~70% suggested that 30% of the patients would have a false-positive COVID-19 test. False-positive results can lead to over-treatment, but false-negative results will prevent patients from receiving essential treatment services in addition to increasing the risk of disease spread to the community. However, these predictions depend on estimates of prevalence. Since prevalence is often highly variable, no meaningful information can be obtained by combining these values. For example, our study estimated a prevalence of olfactory dysfunction ranging from 6% to 84%, whereas for a given diagnostic test neither sensitivity nor specificity will be

affected by the prevalence. Therefore, the importance of sensitivity and specificity would need to be higher for these measures to improve diagnostic accuracy [45,46].

In a previous meta-analysis on the prevalence of olfactory or taste dysfunction in patients with COVID-19, the correlation between self-reported olfactory function and objective measures was generally poor, which may have caused the significant heterogeneity in the summed prevalence. The study classified the University of Pennsylvania Smell Identification Test, sniffin' sticks, and the questionnaire or reporting tool developed by the AAO-HNS as validated instruments [21]. In our study, the same classification was applied to the subgroup analysis, which showed that OTD identified with validated instruments was significantly sensitive (~80%) and specific (> 90%). These results were consistent with the more accurate diagnostic ability of these instruments than of self-reporting in the diagnosis of olfactory disorder [47]. While our results suggest that a validated tool for OT can be used as a screening test, this subgroup analysis included involved only three studies and they were of high heterogeneity. Thus, prospective studies using validated measurement tools for a large number of patients are needed to support our recommendations.

In addition, it has been reported that age, severity of the COVID-19 (mild to moderate or severe), and even ethnicity affects the clinical symptoms of COVID-19 [48-50]. We tried to evaluate the effect of these factors on the olfactory related symptoms. However, since the enrolled studies were limited, the subgroup analysis related to age and ethnicity could not be conducted. On the other hand, regarding the severity, the severe subgroup tended to be less sensitive but more specific than mild to moderate subgroup in the OTD (37% vs 60%; 98% vs 89%). It was recently reported that olfaction related symptoms may not be identified or could be neglected in COVID-19 patients with more severe respiratory symptoms (higher

false negative) [50]. In the view of diagnostic accuracy, low sensitivity can be interpreted as the result of high false negative. This tendency would be partially consistent with the recent study that the overall prevalence was 31% in patients with severe symptoms, lower than 67% in mild-to-moderate symptomatic home-isolated patients. In addition, specificity tend to be inversely related to the sensitivity generally.

Primary physicians and otolaryngologists are likely to be the first clinicians to encounter patients with symptoms suggesting COVID-19 or who are mildly symptomatic. They should therefore be aware of the predictive value of other common symptoms. However, in our study, nasal symptoms, sore throat, and other otorhinolaryngologic symptoms were of no diagnostic value (low sensitivity and specificity of ~20% and ~80%, respectively) for COVID-19 and were not significantly associated with its prevalence. These findings are consistent, and support those of previous reports showing that, unlike other upper respiratory infections, COVID-19 is likely to present with olfactory disorder in the absence of other nasal symptoms. This finding suggests direct viral damage to the chemosensory system [5] and is consistent with both the neuro-invasive tendency of the SARS-CoV-2 virus and the ability of olfactory nerve cells to act as a gateway for neuronal invasion [16].

None of the general non-respiratory (fatigue, fever, headache, diarrhea, and myalgia) or respiratory (cough and dyspnea) symptoms were of high diagnostic accuracy (low sensitivity of 20–60% and moderate specificity of 40–80%). However, non-respiratory symptoms, including diarrhea, fatigue, fever, and myalgia, were significantly associated with COVID-19 positivity, unlike respiratory symptoms such as cough and dyspnea. Possible reasons for this are as follows. Firstly, most of the enrolled studies were retrospective or cross-sectional with self-reporting questionnaires [21]. Thus, recall and selection bias may have led to the over-

presentation of patients with atypical (non-respiratory) symptoms [22]. Secondly, the enrolled studies were comparative and included all patients with upper respiratory tract infection and RT-PCR tests. Accordingly, respiratory symptoms would have been common among the enrolled patients regardless of their COVID-19 status, rather than more common in the COVID-19-positive group. Thirdly, most of the enrolled patients had mild to moderate symptoms, whereas dyspnea, as a marker of more severe COVID-19 disease, might not have been captured in the surveyed studies [5,22]. Therefore, these results are relevant for differentiating COVID-19 from other respiratory infections, not from a healthy condition, and should therefore be interpreted with caution.

Although the diagnosis based the symptoms or signs would be difficult and of low diagnostic value compare with the current reference test, this study is the first meta-analysis to comprehend the clinical meanings of otolaryngological and general symptoms in view of the primary physicians and otolaryngologists to be on the verge of the front-line in the era of COVID-19. In particular, considering the respiratory sampling for PCR requires the personal preventive equipment and is practically impossible in the primary clinics, such knowledge could be helpful to make the presumptive questionnaire for screening and prevent the clinicians from contacting the patients in person. Based on our results, OTD showed higher diagnostic value compared to other otolaryngological and general symptoms among the patients with upper respiratory symptoms. Also, compared to non-validated instruments, validated olfactory and/or taste questionnaires or tools had a clinically high diagnostic accuracy.

Our meta-analysis had several limitations. First, due to the significant heterogeneity of the data pooled in this study, a random effects model and subgroup analysis had to be used. The

source of this heterogeneity was likely to be the wide range (6%–84%) of the reported prevalence of olfactory dysfunction [21]. In addition, RT-PCR from nasopharyngeal swabs is the main diagnostic test for COVID-19 but, as demonstrated in the present study, its sensitivity is only 56–83% [2], which may lead to misclassification and diagnostic bias and thus to a heterogeneity similar to that of prevalence [6]. Second, cross-sectional or retrospective studies have inherent limitations. Together, these two factors may have contributed to an under- or over-estimation of the actual prevalence. A third limitation was the variability of the tools used to assess olfactory and taste dysfunction, as most were self-reporting olfactory and gustatory dysfunction questionnaires, whose weaknesses are well-recognized [21].

Conclusions

Considering the limited accessibility of medical resources, including RT-PCR tests, during the COVID-19 pandemic, screening for OTD or OD may be a valuable tool among patients with influenza-like symptoms. Compared to non-validated instruments, validated questionnaires or tools had a clinically high diagnostic accuracy. Prospective studies with larger numbers are needed to confirm our findings.

Reference

1. Joffily L, Ungierowicz A, David AG, Melo B, Brito CLT, Mello L, et al. The close relationship between sudden loss of smell and COVID-19. Braz J Otorhinolaryngol. 2020 Sep-Oct;86(5):632-8.
2. Zayet S, Klopfenstein T, Mercier J, Kadiane-Oussou NdJ, Lan Cheong Wah L, Royer P-Y, et al. Contribution of anosmia and dysgeusia for diagnostic of COVID-19 in outpatients. Infection. 2020:1-5.
3. Yan CH, Faraji F, Prajapati DP, Boone CE, DeConde AS. Association of chemosensory dysfunction and COVID-19 in patients presenting with influenza-like symptoms. Int Forum Allergy Rhinol. 2020 Jul;10(7):806-13.
4. Sayin I, Yazici ZM. Taste and Smell Impairment in SARS-CoV-2 Recovers Early and Spontaneously: Experimental Data Strongly Linked to Clinical Data. ACS Chem Neurosci. 2020 Jul 15;11(14):2031-3.
5. Roland LT, Gurrola JG, 2nd, Loftus PA, Cheung SW, Chang JL. Smell and taste symptom-based predictive model for COVID-19 diagnosis. Int Forum Allergy Rhinol. 2020 Jul;10(7):832-8.
6. Bénézit F, Le Turnier P, Declerck C, Paillé C, Revest M, Dubée V, et al. Utility of hyposmia and hypogeusia for the diagnosis of COVID-19. Lancet Infect Dis. 2020 Sep;20(9):1014-5.
7. Park JH, Jang W, Kim SW, Lee J, Lim YS, Cho CG, et al. The Clinical Manifestations and Chest Computed Tomography Findings of Coronavirus Disease 2019 (COVID-19) Patients in China: A Proportion Meta-Analysis. Clin Exp Otorhinolaryngol. 2020 May;13(2):95-105.

8. Kim SY, Kim DW. Does the Clinical Spectrum of Coronavirus Disease 2019 (COVID-19) Show Regional Differences? Clin Exp Otorhinolaryngol. 2020 May;13(2):83-4.

9. Menni C, Valdes A, Freydin MB, Ganesh S, El-Sayed Moustafa J, Visconti A, et al. Loss of smell and taste in combination with other symptoms is a strong predictor of COVID-19 infection. 2020:2020.04.05.20048421.

10. Beltrán-Corbellini Á, Chico-García JL, Martínez-Poles J, Rodríguez-Jorge F, Natera-Villalba E, Gómez-Corral J, et al. Acute-onset smell and taste disorders in the context of COVID-19: a pilot multicentre polymerase chain reaction based case-control study. Eur J Neurol. 2020 Apr 22.

11. Just J, Puth M-T, Regenold F, Weckbecker K, Bleckwenn M. Distinguishing between COVID-19 and the common cold in a primary care setting - comparison of patients with positive and negative SARS-CoV-2 PCR results. 2020:2020.04.27.20081877.

12. Boudjema S, Finance J, Coulibaly F, Meddeb L, Tissot-Dupont H, Michel M, et al. Olfactory and gustative disorders for the diagnosis of COVID-19. Travel medicine and infectious disease. 2020 Sep-Oct;37:101875-.

13. Clemency BM, Varughese R, Scheafer DK, Ludwig B, Welch JV, McCormack RF, et al. Symptom Criteria for COVID-19 Testing of Heath Care Workers. Acad Emerg Med. 2020 Jun;27(6):469-74.

14. Dixon BE, Wools-Kaloustian K, Fadel WF, Duszynski TJ, Yiannoutsos C, Halverson PK, et al. Symptoms and symptom clusters associated with SARS-CoV-2 infection in community-based populations: Results from a statewide epidemiological study. medRxiv. 2020 Oct 22.

15. Haehner A, Draf J, Dräger S, de With K, Hummel T. Predictive Value of Sudden Olfactory Loss in the Diagnosis of COVID-19. ORL; journal for oto-rhino-laryngology and its related specialties. 2020;82(4):175-80.

16. Hornuss D, Lange B, Schröter N, Rieg S, Kern WV, Wagner D. Anosmia in COVID-19 patients. Clin Microbiol Infect. 2020 Oct;26(10):1426-7.

17. Jiang J, Wan R, Pan P, Hu C, Zhou R, Yin Y, et al. Comparison of Clinical, Laboratory and Radiological Characteristics Between COVID-19 and Adenovirus Pneumonia: A Retrospective Study. Infection and Drug Resistance. 2020;13:3401-8.

18. Menni C, Sudre CH, Steves CJ, Ourselin S, Spector TD. Quantifying additional COVID-19 symptoms will save lives. Lancet. 2020 Jun 20;395(10241):e107-e8.

19. Menni C, Valdes AM, Freidin MB, Sudre CH, Nguyen LH, Drew DA, et al. Real-time tracking of self-reported symptoms to predict potential COVID-19. Nat Med. 2020 Jul;26(7):1037-40.

20. Moein ST, Hashemian SM, Mansourafshar B, Khorram-Tousi A, Tabarsi P, Doty RL. Smell dysfunction: a biomarker for COVID-19. Int Forum Allergy Rhinol. 2020 Aug;10(8):944-50.

21. Tong JY, Wong A, Zhu D, Fastenberg JH, Tham T. The Prevalence of Olfactory and Gustatory Dysfunction in COVID-19 Patients: A Systematic Review and Meta-analysis. Otolaryngol Head Neck Surg. 2020 Jul;163(1):3-11.

22. Tostmann A, Bradley J, Bousema T, Yiek WK, Holwerda M, Bleeker-Rovers C, et al. Strong associations and moderate predictive value of early symptoms for SARS-CoV-2 test positivity among healthcare workers, the Netherlands, March 2020. Euro Surveill. 2020 Apr;25(16).

23. Wee LE, Chan YFZ, Teo NWY, Cherng BPZ, Thien SY, Wong HM, et al. The role of self-reported olfactory and gustatory dysfunction as a screening criterion for suspected COVID-19. Eur Arch Otorhinolaryngol. 2020 Aug;277(8):2389-90.

24. Zhao D, Yao F, Wang L, Zheng L, Gao Y, Ye J, et al. A Comparative Study on the Clinical Features of Coronavirus 2019 (COVID-19) Pneumonia With Other Pneumonias. Clin Infect Dis. 2020 Jul 28;71(15):756-61.

25. Zhu W, Xie K, Lu H, Xu L, Zhou S, Fang S. Initial clinical features of suspected coronavirus disease 2019 in two emergency departments outside of Hubei, China. J Med Virol. 2020 Sep;92(9):1525-32.

26. Cheng Z, Lu Y, Cao Q, Qin L, Pan Z, Yan F, et al. Clinical Features and Chest CT Manifestations of Coronavirus Disease 2019 (COVID-19) in a Single-Center Study in Shanghai, China. AJR Am J Roentgenol. 2020 Jul;215(1):121-6.

27. Sun Y, Koh V, Marimuthu K, Ng OT, Young B, Vasoo S, et al. Epidemiological and Clinical Predictors of COVID-19. Clin Infect Dis. 2020 Jul 28;71(15):786-92.

28. Tolia VM, Chan TC, Castillo EM. Preliminary Results of Initial Testing for Coronavirus (COVID-19) in the Emergency Department. West J Emerg Med. 2020 Mar 27;21(3):503-6.

29. Peng L, Liu KY, Xue F, Miao YF, Tu PA, Zhou C. Improved Early Recognition of Coronavirus Disease-2019 (COVID-19): Single-Center Data from a Shanghai Screening Hospital. Arch Iran Med. 2020 Apr 1;23(4):272-6.

30. Nobel YR, Phipps M, Zucker J, Lebwohl B, Wang TC, Sobieszczyk ME, et al. Gastrointestinal Symptoms and Coronavirus Disease 2019: A Case-Control Study From the United States. Gastroenterology. 2020 Jul;159(1):373-5.e2.

31. Chen X, Tang Y, Mo Y, Li S, Lin D, Yang Z, et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. Eur Radiol. 2020 Sep;30(9):4893-902.

32. Boscolo-Rizzo P, Borsetto D, Spinato G, Fabbris C, Menegaldo A, Gaudioso 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. 2020 Sep;277(9):2637-40.

33. Carignan A, Valiquette L, Grenier C, Musonera JB, Nkengurutse D, Marcil-Héguy A, et al. Anosmia and dysgeusia associated with SARS-CoV-2 infection: an age-matched case-control study. Cmaj. 2020 Jun 29;192(26):E702-e7.

34. Chung TW, Sridhar S, Zhang AJ, Chan KH, Li HL, Wong FK, et al. Olfactory Dysfunction in Coronavirus Disease 2019 Patients: Observational Cohort Study and Systematic Review. Open Forum Infect Dis. 2020 Jun;7(6):ofaa199.

35. Hoang MP, Kanjanaumporn J, Aeumjaturapat S, Chusakul S, Seresirikachorn K, Snidvongs K. Olfactory and gustatory dysfunctions in COVID-19 patients: A systematic review and meta-analysis. Asian Pac J Allergy Immunol. 2020 Sep;38(3):162-9.

36. Rojas-Lechuga MJ, Izquierdo-Domínguez A, Chiesa-Estomba C, Calvo-Henríquez C, Villarreal IM, Cuesta-Chasco G, et al. Chemosensory dysfunction in COVID-19 out-patients. Eur Arch Otorhinolaryngol. 2020 Aug 25:1-8.

37. Feng C, Huang Z, Wang L, Chen X, Zhai Y, Zhu F, et al. A Novel Triage Tool of Artificial Intelligence Assisted Diagnosis Aid System for Suspected COVID-19 pneumonia In Fever Clinics. medRxiv. 2020:2020.03.19.20039099.

38. Song C-Y, Xu J, He J-Q, Lu Y-Q. COVID-19 early warning score: a multi-parameter

screening tool to identify highly suspected patients. medRxiv. 2020:2020.03.05.20031906.

39. Yang Z, Lin D, Chen X, Qiu J, Li S, Huang R, et al. Distinguishing COVID-19 from influenza pneumonia in the early stage through CT imaging and clinical features. medRxiv. 2020:2020.04.17.20061242.

40. Ai J-W, Zhang H-C, Xu T, Wu J, Zhu M, Yu Y-Q, et al. Optimizing diagnostic strategy for novel coronavirus pneumonia, a multi-center study in Eastern China. medRxiv. 2020:2020.02.13.20022673.

41. Rentsch CT, Kidwai-Khan F, Tate JP, Park LS, King JT, Skanderson M, et al. Covid-19 Testing, Hospital Admission, and Intensive Care Among 2,026,227 United States Veterans Aged 54-75 Years. medRxiv : the preprint server for health sciences. 2020:2020.04.09.20059964.

42. Liang Y, Liang J, Zhou Q, Li X, Lin F, Deng Z, et al. Prevalence and clinical features of 2019 novel coronavirus disease (COVID-19) in the Fever Clinic of a teaching hospital in Beijing: a single-center, retrospective study. medRxiv. 2020:2020.02.25.20027763.

43. Kim DH, Kim Y, Kim SW, Hwang SH. Use of narrowband imaging for the diagnosis and screening of laryngeal cancer: A systematic review and meta-analysis. Head Neck. 2020 Sep;42(9):2635-43.

44. Vijayakumar V, Reghunathan D, Edacheriyan B, Mukundan A. Role of Toluidine Blue Staining in Suspicious Lesions of Oral Cavity and Oropharynx. Indian J Otolaryngol Head Neck Surg. 2019 Oct;71(Suppl 1):142-6.

45. Trikalinos TA, Balion CM, Coleman CI, Griffith L, Santaguida PL, Vandermeer B, et

al. AHRQ Methods for Effective Health Care Meta-Analysis of Test Performance When There Is a "Gold Standard". In: Chang SM, Matchar DB, Smetana GW, Umscheid CA, editors. Methods Guide for Medical Test Reviews. Rockville (MD): Agency for Healthcare Research and Quality (US); 2012.

46. Trikalinos TA, Balion CM, Coleman CI, Griffith L, Santaguida PL, Vandermeer B, et al. Chapter 8: meta-analysis of test performance when there is a "gold standard". J Gen Intern Med. 2012 Jun;27 Suppl 1:S56-66.

47. Yang J, Pinto JM. The Epidemiology of Olfactory Disorders. Curr Otorhinolaryngol Rep. 2016 May;4(2):130-41.

48. Hall A, Frauenfelder C, Butler C, Coyle P, Hopkins C. Paediatric olfactory dysfunction: a chance to detect COVID-19? Arch Dis Child. 2020 Aug 20.

49. von Bartheld CS, Hagen MM, Butowt R. Prevalence of Chemosensory Dysfunction in COVID-19 Patients: A Systematic Review and Meta-analysis Reveals Significant Ethnic Differences. ACS Chem Neurosci. 2020 Oct 7;11(19):2944-61.

50. Borsetto D, Hopkins C, Philips V, Obholzer R, Tirelli G, Polesel J, 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. 2020 Oct 1;58(5):430-6.

Table 1. Comparison of the diagnosis accuracy in other ENT symptoms

Subgroup	Study (n)	DOR [95% CIs]	Sensitivity [95% CIs]/I^2	Specificity [95% CIs]	AUC	NPV	PPV
Other ENT symptom							
Nasal discharge	14	1.02 [0.60; 1.74]/ I^2 = 86.8%	0.18 [0.10; 0.31]/ I^2 = 96.1%	0.81 [0.69; 0.89]/ I^2 = 99.0%	0.46	0.75 [0.59; 0.87]/ I^2= 99.0%	0.21 [0.08; 0.44]/ I^2 = 98.2%
Nasal obstruction	8	0.98 [0.76; 1.26]/ I^2 = 0.0%	0.20 [0.11; 0.35] / I^2 = 92.7%	0.75 [0.65; 0.83]/ I^2 = 92.0%	0.55	0.62 [0.39; 0.81]/ I^2 = 98.1%	0.31 [0.12; 0.60]/ I^2 = 96.3%
Sore throat	16	0.66 [0.38; 1.15] / I^2 = 91.0%	0.24 [0.16; 0.33]/ I^2 = 91.2%	0.68 [0.58; 0.77]/ I^2 = 98.3%	0.41	0.73 [0.56; 0.86]/ I^2 = 98.9%	0.18 [0.090; 0.33]/ I^2 = 97.5%

DOR; diagnostic odds ratio, AUC; area under the curve, NPV; negative predictive value, PPV; positive predictive value

Table 2. Comparison of the diagnosis accuracy in General symptom

Subgroup	Study (n)	DOR [95% CIs]	Sensitivity [95% CIs]/I^2	Specificity [95% CIs]	AUC	NPV	PPV
General symptom							
Cough	24	0.98 [0.73; 1.32]/ I^2 = 91.5%	0.59 [0.54; 0.64] / I^2 = 89.7%	0.39 [0.30; 0.48]/ I^2 = 99.2%	0.5	0.69 [0.55; 0.80]/ I^2 = 99.2%	0.29 [0.20; 0.39]/ I^2 = 98.8%
Diarrhea	19	1.34 [1.09; 1.66]/ I^2 = 67.4%	0.17 [0.11; 0.25]/ I^2 = 97.2%	0.85 [0.78; 0.90]/ I^2 = 98.9%	0.51	0.71 [0.60; 0.80]/ I^2 = 99.3%	0.31 [0.21; 0.43]/ I^2 = 97.1%
Dyspnea	17	1.12 [0.78; 1.61]/ I^2 = 87.6%	0.18 [0.13; 0.24]/ I^2 = 92.0%	0.84 [0.76; 0.90]/ I^2 = 99.0%	0.39	0.74 [0.59; 0.85] / I^2 = 99.6%	0.25 [0.13; 0.42]/ I^2 = 98.4%
Fatigue	12	1.67 [1.20; 2.34]/ I^2 = 89.8%	0.35 [0.25; 0.46]/ I^2 = 96.9%;	0.76 [0.61; 0.86]/ I^2 = 99.6%	0.52	0.75 [0.61; 0.84]/ I^2 = 99.3%	0.35 [0.19; 0.55]/ I^2 = 99.2%
Fever	22	2.22 [1.43; 3.44]/ I^2 = 95.8%	0.60 [0.47; 0.73]/ I^2 = 98.5%	0.55 [0.38; 0.71]/ I^2 = 99.7%	0.62	0.81 [0.70; 0.88]/ I^2 = 99.4%	0.33 [0.25; 0.43]/ I^2 = 98.1%
Headache	15	1.48 [0.98; 2.22]/ I^2 = 77.2%	0.36 [0.21; 0.53] / I^2 = 96.5%	0.74 [0.58; 0.86]/ I^2 = 99.1%	0.57	0.74 [0.55; 0.87]/ I^2 = 98.6%	0.35 [0.20; 0.53] / I^2 = 97.4%
Myalgia	12	2.09 [1.20; 3.64]/ I^2 = 90.1%	0.48 [0.33; 0.64]/ I^2 = 95.3%	0.68 [0.55; 0.79]/ I^2 = 98.3%	0.61	0.82 [0.67; 0.91]/ I^2 = 98.3%	0.29 [0.19; 0.42]/ I^2 = 95.7%

DOR; diagnostic odds ratio, AUC; area under the curve, NPV; negative predictive value, PPV; positive predictive value

Supplementary Table 1. Search terms and queries of the database.

Supplementary Table 2. Study characteristics.

Supplementary Table 3. Methodological qualities and bias assessment of all included studies.

Figure Legends

Figure 1. Summary of the search strategy.

Figure 2. Forest plot of the diagnostic odds ratios of the included studies. Olfactory and/or taste disorder (A) and only olfactory disorder (B).

Figure 3. Area under the summary receiver operating characteristic of the included studies. Olfactory and/or taste disorder (A) and only olfactory disorder (B).

Supplementary Figure 1. Forest plots of the sensitivity (A), specificity (B), negative predictive values (C), and positive predictive values (D) of the olfactory and/or taste disorder studies.

Supplementary Figure 2. Forest plots of the sensitivity (A), specificity (B), negative predictive values (C), and positive predictive values (D) of the only olfactory disorder studies.

The English in this document has been checked by at least two professional editors, both native speakers of English. For a certificate, please see:

http://www.textcheck.com/certificate/3B9rRw

Figure 1

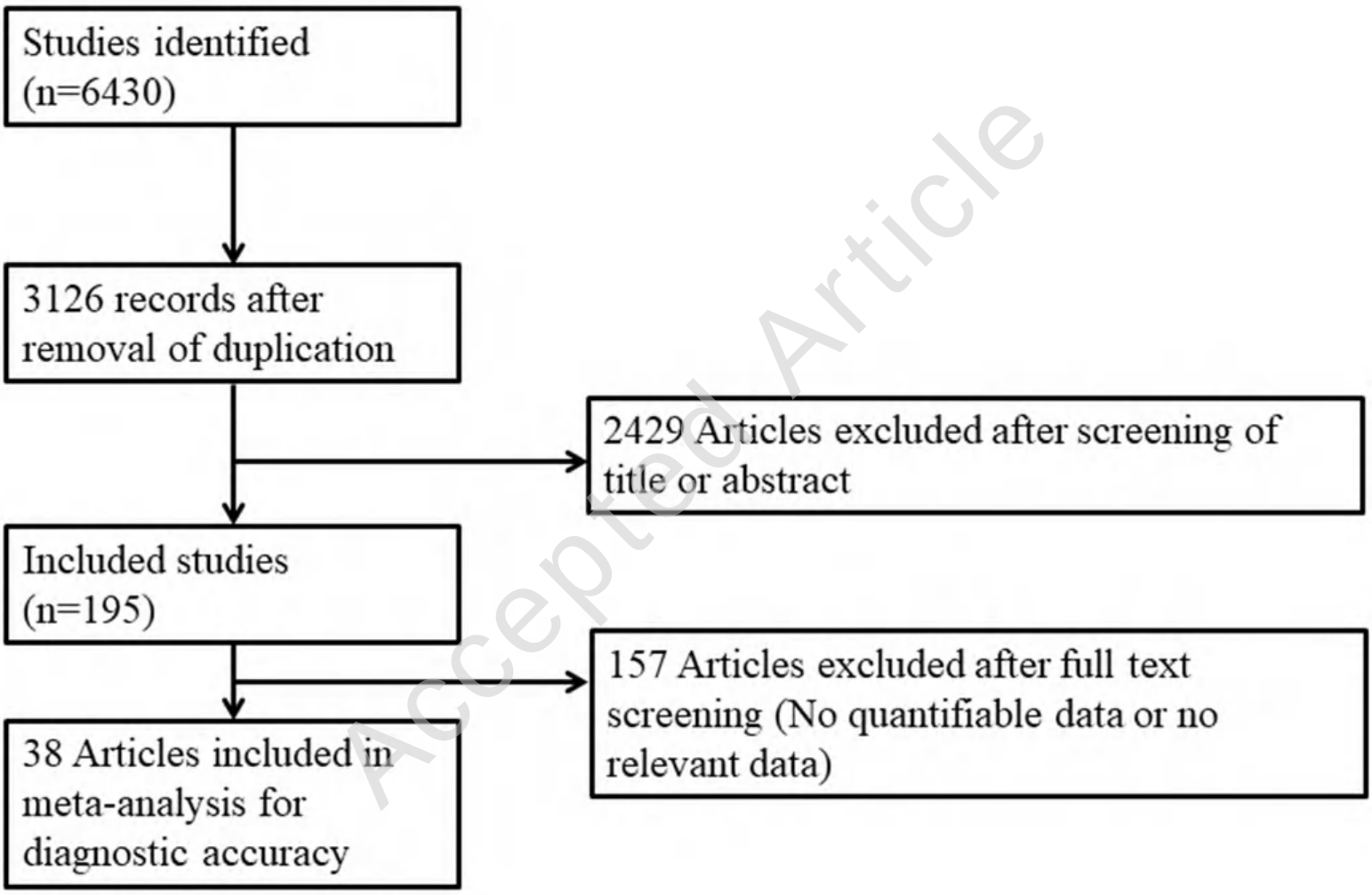
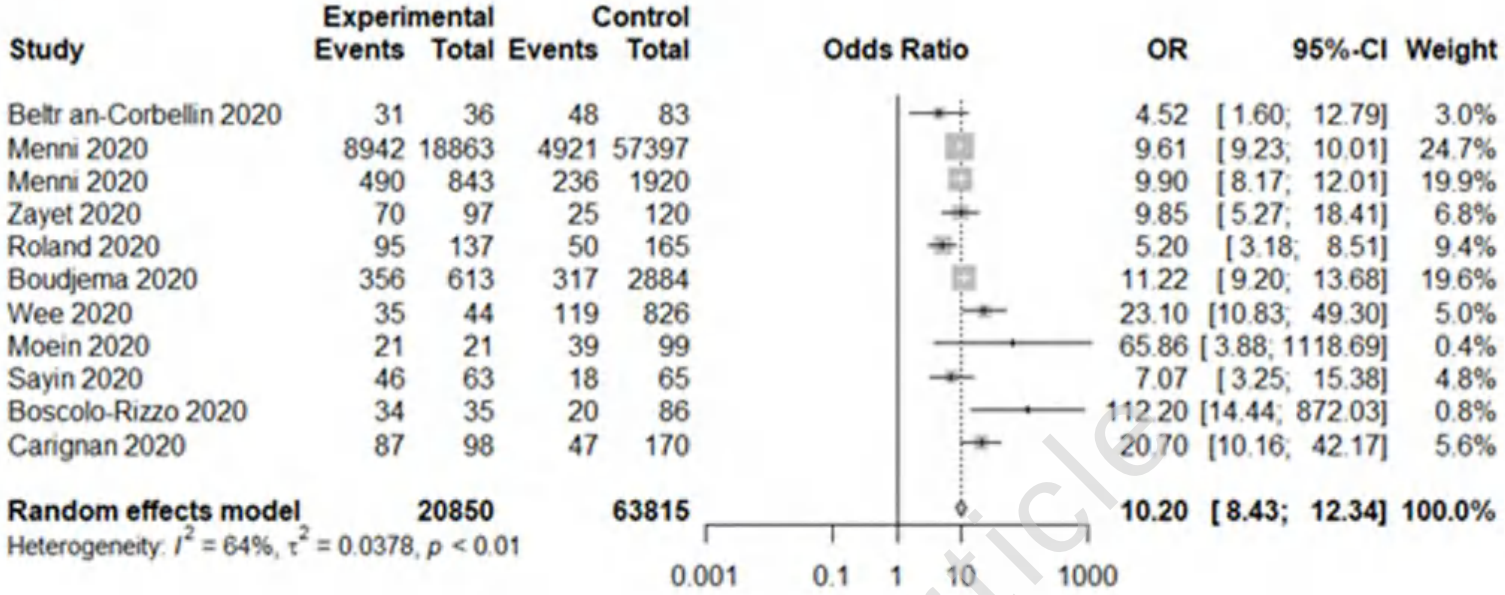


Figure 2

A



B

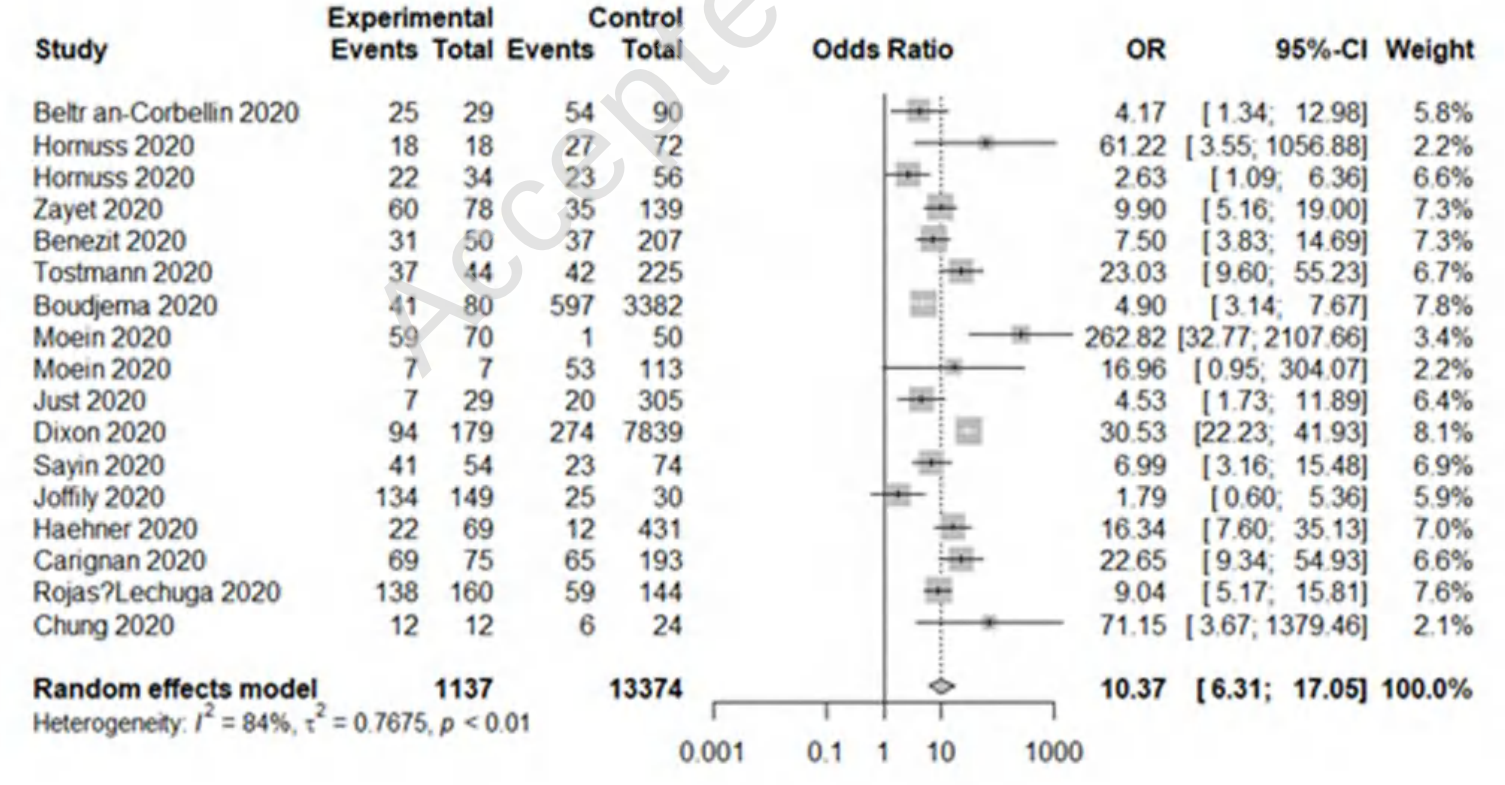
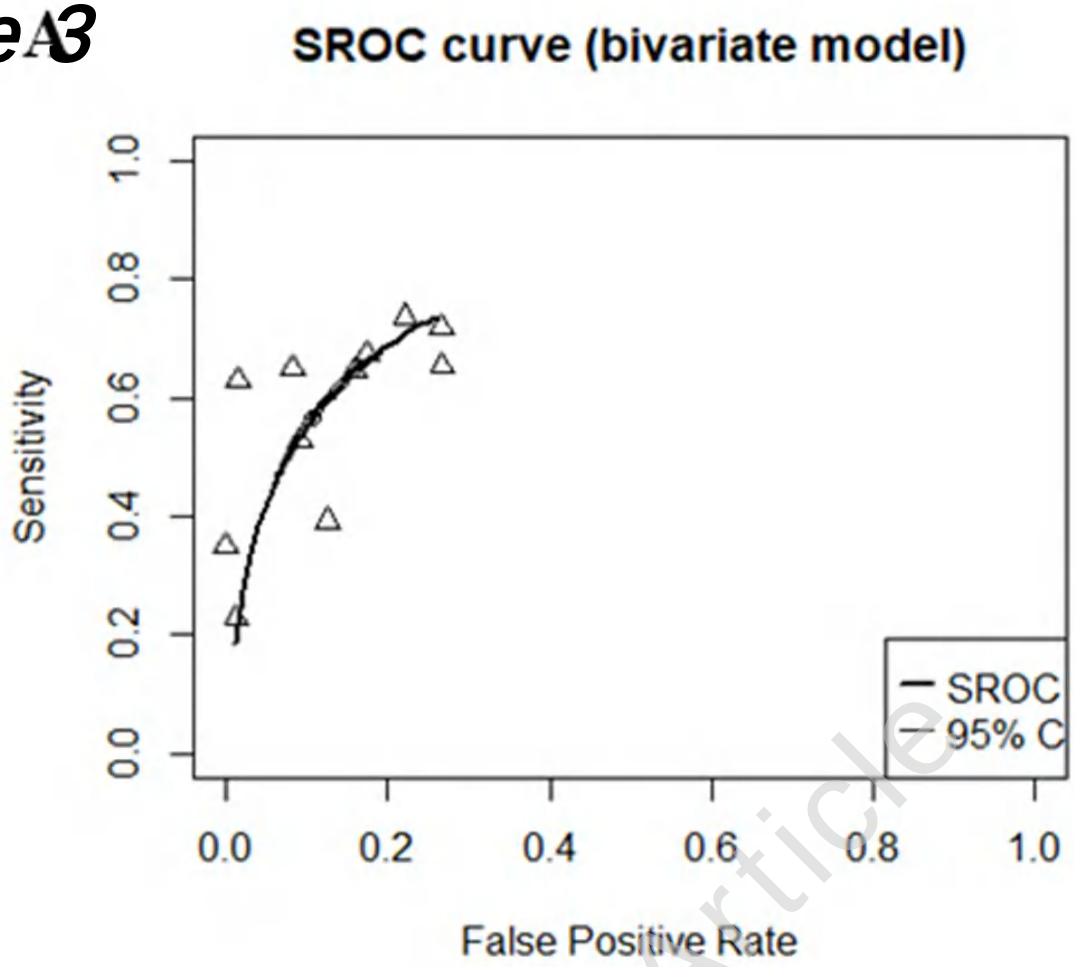
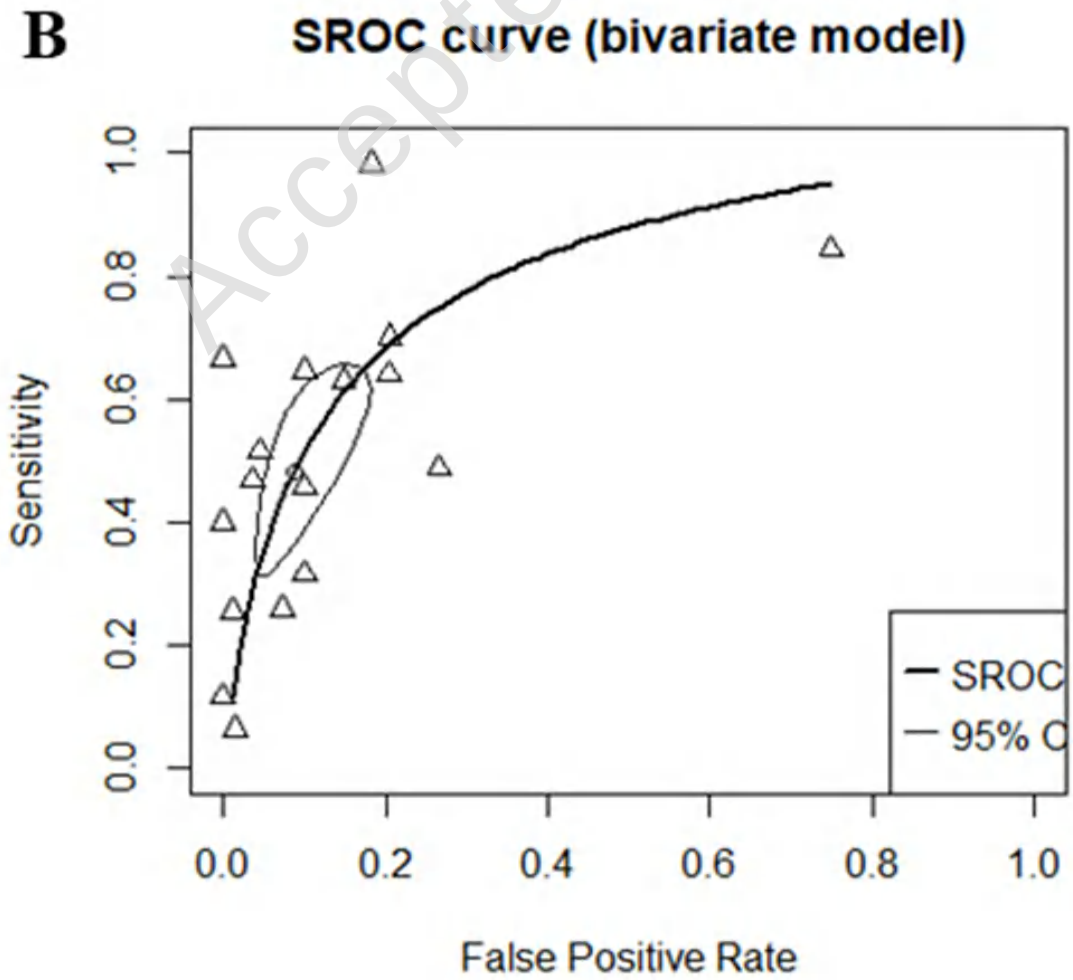
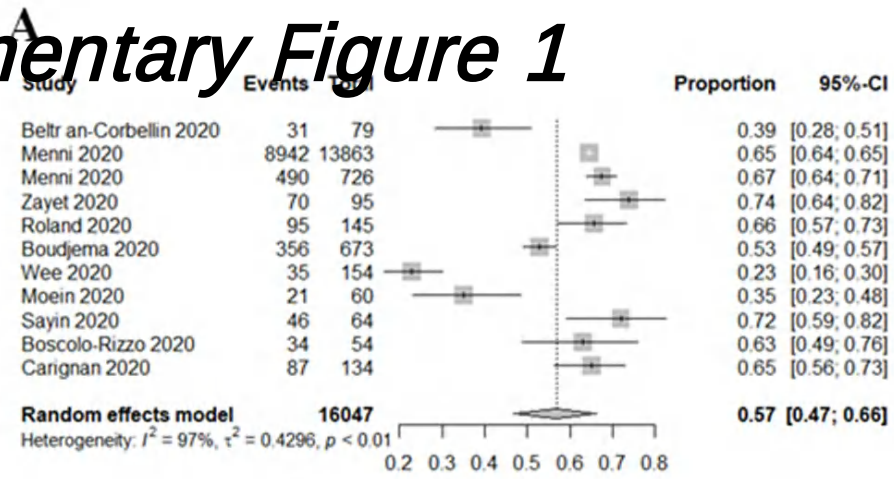


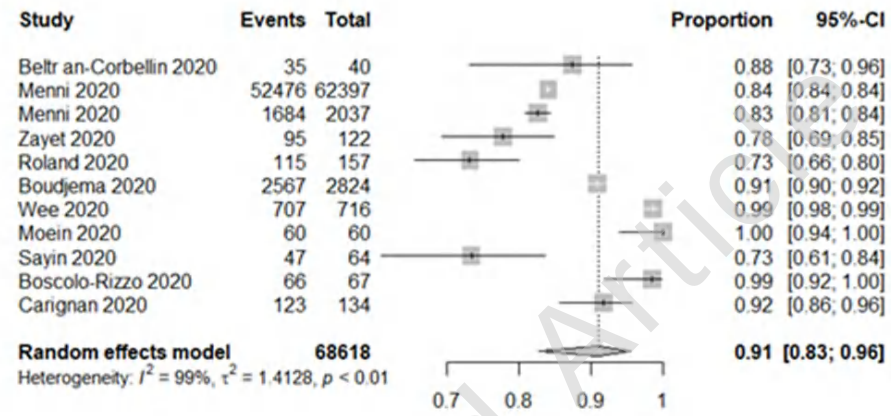
Figure 3

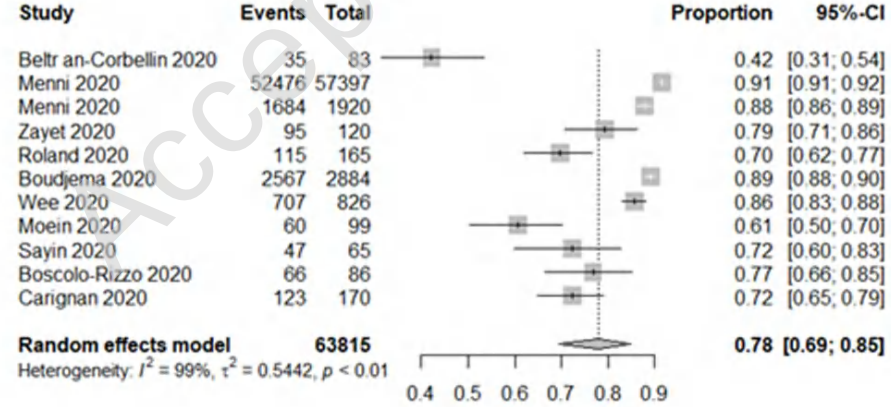


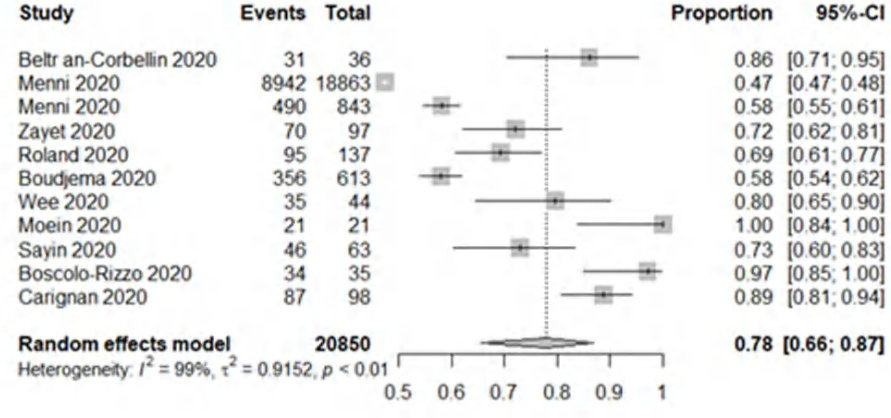


Supplementary Figure 1









Supplementary Figure 2