BMJ Open

The Diagnostic Accuracy of Subjective Dyspnea in Detecting Hypoxemia Among Outpatients with COVID-19: a retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-046282
Article Type:	Original research
Date Submitted by the Author:	28-Oct-2020
Complete List of Authors:	Berezin, Linor; University of Toronto Faculty of Medicine Zhabokritsky, Alice; University of Toronto Faculty of Medicine Andany, Nisha; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto, Faculty of Medicine Chan, Adrienne ; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Estrada-Codecido, Jose; Sunnybrook Health Sciences Centre, Division of Infectious Diseases Gershon, Andrea; Sunnybrook Health Sciences Centre, Division of Respirology; University of Toronto Faculty of Medicine Lam, Philip; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Leis, Jerome; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine MacPhee, Scott; Sunnybrook Health Sciences Centre, Department of Nursing Mubareka, Samira; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto, Department of Laboratory Medicine and Pathology Simor, Andrew; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Daneman, Nick; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine, Medicine
Keywords:	COVID-19, INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The Diagnostic Accuracy of Subjective Dyspnea in Detecting Hypoxemia Among Outpatients with COVID-19: a retrospective cohort study

Author list:

Linor Berezin[a], Alice Zhabokritsky[a], Nisha Andany[a,b], Adrienne K. Chan[a,b], Jose Estrada-Codecido[b], Andrea Gershon[a,c], Philip W. Lam[a,b], Jerome A. Leis[a,b], Scott MacPhee[d], Samira Mubareka[b,e], Andrew E. Simor[a,b], Nick Daneman[a,b].

Author Affiliations:

[a] Faculty of Medicine, University of Toronto, Toronto, Canada

[b] Division of Infectious Diseases, Sunnybrook Health Sciences Centre, University of Toronto

[c] Division of Respirology, Sunnybrook Health Sciences Centre, University of Toronto

[d] Department of Nursing, Sunnybrook Health Sciences Centre, University of Toronto

[e] Department of Laboratory Medicine and Pathology, University of Toronto, Toronto, Canada

Corresponding author:

Nick Daneman, MD, FRCPC, MSc

Division of Infectious Diseases, Sunnybrook Health Sciences Centre, University of Toronto,

2075 Bayview Avenue, Toronto, ON, M4N 2M5, Canada

nick.daneman@sunnybrook.ca

Key Words: COVID-19, Dyspnea, hypoxemia, oxygen saturation, monitoring, virtual care

Word Count: 2549

Author Contributions: The authors all stand behind the conclusions of this manuscript, agree to be accountable for all aspects of the work, and support its publication. All authors contributed to the study conception and designed the protocol. All authors contributed to the manuscript preparation and have given approval for its submission.

Conflicts of Interest: The authors report no financial or other conflicts of interest.

Abstract:

Objectives: The majority of patients with mild-to-moderate COVID-19 can be managed using virtual care. Dyspnea is challenging to assess remotely, and the accuracy of subjective dyspnea measures in capturing hypoxemia have not been formally evaluated for COVID-19. We explored the accuracy of subjective dyspnea in diagnosing hypoxemia in COVID-19 patients.

Methods: This is a retrospective cohort study of consecutive outpatients with COVID-19 who met criteria for home oxygen saturation monitoring at a university-affiliated acute care hospital in Toronto, Canada from April 3, 2020 to September 13, 2020. Dyspnea measures were treated as diagnostic tests, and we determined their sensitivity (SN), specificity (SP), negative/positive predictive value (NPV/PPV), and positive/negative likelihood ratios (+LR/-LR) for detecting hypoxemia. In the primary analysis, hypoxemia was defined by oxygen saturation <95%; the diagnostic accuracy of subjective dyspnea was also assessed across a range of oxygen saturation cutoffs from 92% to 97%.

Results: During the study period 89/501 (17.8%) of patients met criteria for home oxygen saturation monitoring, and of these 17/89 (19.1%) were diagnosed with hypoxemia. The presence/absence of dyspnea had limited accuracy for diagnosing hypoxemia, with SN 47% (95%CI 24-72%), SP 80% (68%-88%), NPV 86% (75%-93%), PPV 36% (18%-59%), +LR 2.4 (1.2-4.7), -LR 0.7 (0.4-1.1). The SN of dyspnea was 50% (95%CI 19-81) when a cutoff of ≤92% was used to define hypoxemia. An mMRC dyspnea score >1 (SP 98%, 95%CI 88%-100%), Roth Maximal Count <12 (SP 100%, 95%CI 75-100%), and Roth Counting time < 8 seconds (SP 93%, 95%CI 66%-100%) had high SP that could be used to rule in hypoxemia, but displayed low SN (≤50%).

Conclusions: Subjective dyspnea measures have inadequate accuracy for ruling out hypoxemia in high-risk patients with COVID-19. Safe home management of patients with COVID-19 should incorporate home oxygenation saturation monitoring.

Strengths and limitations of the study:

- This is the first study to evaluate the diagnostic accuracy of subjective dyspnea in detecting hypoxemia in the setting of COVID-19.
- We provide evidence justifying the need for home oxygen saturation monitoring that will inform the safe home management of outpatients with COVID-19
- This study was limited to patients who were considered high risk for severe COVID-19, which limited the sample size and it is possible that the diagnostic test characteristics might differ in younger, healthier patients.
- The data collected for this study were from single patient assessments and did not assess whether changes in dyspnea correlate with changes in SpO2 over time.

Introduction:

As of October 23, 2020, there have been more than 41 million laboratory-confirmed novel coronavirus disease 2019 (COVID-19) cases and 1.1 million deaths documented globally.[1] The spectrum of disease of COVID-19 ranges from asymptomatic or mild symptoms, to severe respiratory failure and death.[2] Approximately 20% of patients with COVID-19 experience dyspnea, which is more commonly associated with severe disease.[3] Fatal cases of COVID-19 have higher rates of dyspnea, lower blood oxygen saturation (SpO2), and greater rates of complications such as acute respiratory distress syndrome.[4,5]

In an effort to reduce avoidable hospitalizations, health care contacts, and transmission, most patients with COVID-19 can be managed in the community using virtual healthcare platforms, and transferred to hospital only if they develop progressive respiratory disease.[6] Subjective dyspnea can be assessed remotely using patient interview, and augmented by surrogate measures such as the Roth Score and modified Medical Research Council (mMRC) Dyspnea Scale. However, the accuracy of these measures has not been formally evaluated in the context of COVID-19.[6,7] Of great concern is the risk of false reassurance if patients develop hypoxemia without subjective sensation of dyspnea. “Silent hypoxemia”, or low SpO2 in the absence of dyspnea, has been reported in the setting of COVID-19 and clinicians have speculated that it may be associated with increased out-of-hospital mortality;[8] case reports have described patients presenting to hospital with rapid deterioration and respiratory failure without signs of respiratory distress.[9-11]

Previous studies of the utility of dyspnea measurement in diagnosing hypoxemia in other respiratory conditions such as chronic obstructive pulmonary disease (COPD), congestive heart failure and lung cancer have yielded conflicting results;[12-15] This association has not been studied during the COVID-19 pandemic. Therefore, we sought to determine the diagnostic accuracy of subjective dyspnea measures in diagnosing hypoxemia among a cohort of outpatients with COVID-19.

Methods:

Study Participants

All consecutive patients with laboratory-confirmed COVID-19 followed as outpatients by the Sunnybrook Health Sciences Centre COVID-19 Expansion to Outpatients (COVIDEO) virtual care service from April 3, 2020 to September 13, 2020 were included in this cohort study. The patients were diagnosed based on a positive mid-turbinate or nasopharyngeal swab for COVID-19 RNA detected by real-time polymerase chain reaction. COVIDEO is a virtual care model for monitoring of outpatients with COVID-19 at Sunnybrook Health Sciences Centre, and is the basis of similar programs at other hospitals.[16] Patients were contacted by an infectious diseases physician for assessment and monitoring either by telephone or through the Ontario Telemedicine Network virtual platform.

A portable pulse oximeter was delivered to the homes of high-risk patients as defined by age > 60 years, pregnancy, extensive comorbidities, or presence of cardio-respiratory symptoms, such as chest pain or dyspnea. Patients were instructed to record their SpO2 measurements twice daily throughout their illness. This study was approved by the institutional review board at

Sunnybrook Health Sciences as minimal risk research, using data collected for routine clinical practice, and the requirement for informed consent was waived.

Data Collection

The demographic characteristics, clinical data, measures of subjective dyspnea (presence of shortness of breath, mMRC dyspnea scale score, Roth Score), physical exam findings, and SpO2 readings for study participants were collected from electronic medical records by one investigator (either A.Z or S.Ma.). The subjective dyspnea measures obtained from the patient's first visit with a pulse oximeter were used for analysis.

Predictor Variables

The primary predictor of interest was patient-reported presence of dyspnea. Secondary predictor variables were patient-reported breathing faster at rest, breathing harder than normal, feeling more breathless today than yesterday, as well as dyspnea as measured by the mMRC Dyspnea Scale and the Roth Score.

The mMRC Dyspnea Scale has been studied extensively in a variety of respiratory conditions.[17] It is comprised of five categories that describe the degree of activity limitation due to worsening breathlessness. The participant assigns themselves a score ranging from 0 to 4 based on their perception of which activities result in dyspnea, with higher scores indicating a greater impairment in their ability to perform daily activities.

The Roth Score is a tool for quantifying the severity of dyspnea, in which the patient is asked to count audibly to 30 in their native language, and the maximal count and counting time are recorded. A prior validation study demonstrated a strong positive correlation between pulse oximetry measurement and both counting time ($r = 0.59$; $P < 0.001$) and maximal count ($r = 0.67$; $P < 0.001$) achieved in one breath.[7]

Outcomes

The reference measure was SpO2 as measured by a ChoiceMMed pulse oximeter (model MD300C20). In the primary outcome definition, hypoxemia was considered to be present if SpO2 was < 95% in order to provide sufficient power to estimate the diagnostic test characteristics. Secondary outcomes included hypoxemia cut-offs varying from 92 to 97%. Patients received instructions on correct pulse oximeter use and were told to wait 5 to 10 seconds for readings to calibrate prior to recording the SpO2 measurements.

Statistical Analysis

In the primary analysis, the subjective dyspnea measures were treated as diagnostic tests, and the specificity (SP), sensitivity (SN), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (+LR) and negative likelihood ratio (-LR) were determined in order to evaluate the predictive value in detecting hypoxemia. For the continuous predictors, the test characteristics were provided across a range of different score thresholds.

Diagnostic test characteristics of the primary dyspnea measure were also determined in subgroups stratified based on patient characteristics, including (1) age <60 vs >60 years, (2)

presence vs absence of underlying lung disease, and date from symptom onset (<7 vs ≥7 days). The Wilson method with continuity correction was used to calculate 95% confidence intervals.

A secondary analysis examined the strength of association between the presence of dyspnea and hypoxemia with a χ^2 test or Fisher exact test with dyspnea treated as a binary variable (present or absent). This relationship is represented by a violin plot. In additional analyses, a correlation coefficient was calculated to assess whether there was an association between the participants' Roth Scores and their oxygen saturation measurements. These associations were displayed graphically with a scatter plot. The relationship between the mMRC Dyspnea Scale and hypoxemia was analyzed with a χ^2 test and represented by violin plot.

All analyses were conducted in SAS Statistical Software V.9.3 (Cary, North Carolina, USA). For all statistical analyses, $P < 0.05$ was considered significant.

Sample Size Calculation

The primary test characteristic of interest was the SN of dyspnea as a test for hypoxemia. The sample size was estimated based on a test of single proportion, namely SN. It was estimated that 62 patients would be required in order to estimate a SN with a 10% margin of error and 95% confidence if the true sensitivity was 80%.

Results:

Demographic and Clinical Characteristics of Outpatients with COVID-19

From April 3 to Sept 13, 2020, a total of 501 patients with COVID-19 were followed by COVIDEO. Of these patients, 89 (17.8%) met criteria for home oxygen saturation monitoring. One patient was lost to follow up after provision of the oxygen monitoring device.

Overall, the median age of patients was 52 years (interquartile range (IQR) 38-64 years) and 57 patients (64%) were female (Table 1). The median number of days from symptom onset to clinical assessment was 6 (IQR 3-8). Among these patients, the most common comorbidities were hypertension (36%), obesity (20%), diabetes (17%), asthma (16%), and malignancy (16%). Twenty-nine patients (33%) had no comorbidities. The most common symptoms reported on intake assessment were fatigue/malaise (66%), cough (63%), and myalgias (45%). While the patients were being followed by COVIDEO, 11 (12%) required hospitalization, with a median duration of hospitalization of 3 days (IQR 2.5-7). Five (6%) patients were admitted to the intensive care unit (ICU), and the median length of ICU stay was 6 days (IQR 2-11). One patient was intubated, and no patients died within 30 days of their COVID-19 diagnosis.

Association of Dyspnea Measurements with Detection of Hypoxemia

A total of 17 (19.1%) patients were diagnosed with hypoxemia. Hypoxemia was significantly associated with the presence of dyspnea (p= 0.046), mMRC Dyspnea Scale score over 0 (p= 0.014), over 1 (p= 0.001) and over 2 (p= 0.001) (Table 2; Figure 1). Weak associations were identified between patients' Roth Scores and their oxygen saturation measurements for maximum count (r =0.29; p=0.23) and counting time (r =0.12; p=0.617), respectively.

Diagnostic Accuracy of Dyspnea Measurements in the Detection of Hypoxemia:

The presence or absence of subjective dyspnea had a SN 47% (95%CI 24-72%), SP 80% (68%-88%), NPV 86% (75%-93%), PPV 36% (18%-59%), +LR 2.4 (1.2-4.7), -LR 0.7 (0.4-1.1) for diagnosing hypoxemia (Table 3). The presence of subjective dyspnea had lower SN (25% [16%-37%]), 27% [17%-41%], 40% (23%-59%) for detecting hypoxemia as defined by thresholds of ≤97%, ≤96%, and ≤95%, respectively. At a lower SpO2 threshold of ≤92%, the SN increased only slightly to 50% (19%-81%). The other binary measures of subjective dyspnea, including breathing faster at rest, breathing harder than normal, and feeling more breathless than the day before had lower SN (0% [0%-24%], 0% [0%-24%], and 6.2% [0.3%-32%], respectively), and higher SP (96% [87%-99%], 97% [89%-100%], and 96% [87%-99%], respectively) (Table 3).

mMRC Dyspnea Scale scores were recorded and available for 63 patients (70.8%). An mMRC Dyspnea Scale score of greater than 0 was determined to have a SN of 54% (26%-80%), SP 82% (68%-91%), NPV 87% (74%-95%), PPV 44% (21-69%), -LR 0.6 (0.3-1.0), and +LR 3.0 (1.4-6.5) for the detection of hypoxemia. At higher cutoff values, the SN of the mMRC Dyspnea Scale was reduced to 39% (15%-68%) for scores greater than 1 and 2 and 8% (0.4%-38%) for scores greater than 3. The SP for mMRC Dyspnea Scale scores greater than 1, 2, and 3 at capturing hypoxemia was 98% (88%-100%).

Roth Scores were available for 19 patients (29.7%). The Roth Score had a higher SN at higher cutoff values for counting time. A maximum count of less than 12 had a SN of 25% (1.3%-78%), SP 100% (75%-100%), NPV 83% (58%-96%), PPV 100% (6%-100%), and a -LR of 0.75 (0.4-1.3). The diagnostic test with the highest SN for diagnosing hypoxemia was a Roth score maximum counting time of less than 25 seconds, which still had a SN of only 75% (22%-99%),

and inadequate SP 13% (2.3%-42%), NPV 67% (13%-98%), PPV 19% (5.0%-46%), -LR 1.88 (0.2-16), and +LR 0.87 (0.48-1.6).

The diagnostic accuracy of dyspnea presence in the detection of hypoxemia was most impacted when stratified by the presence of underlying lung disease. In the patients with underlying lung disease, the SN and SP of the presence of dyspnea in detecting hypoxemia was 100% (20%-100%) and 80% (51%-95%), respectively. A lower SN (22% [3.9%-60%]) and high SP (96% [79%-100%]) was observed for patients over 60 years when results were stratified based on age. Stratifying based on days from symptom onset did not impact the diagnostic accuracy of dyspnea in detecting hypoxemia (Table 4).

Discussion:

To our knowledge, this is the first study to evaluate the diagnostic accuracy of subjective dyspnea in detecting hypoxemia in the setting of COVID-19. Self-reported shortness of breath has very limited utility for detecting hypoxemia, with a SN of only 47% and SP of only 80% for detecting SpO2 levels below 95%. Using a lower SpO2 threshold of less than 93% did not meaningfully improve the SN of subjective dyspnea in diagnosing hypoxemia (SN 50%). An mMRC Dyspnea Score exceeding 1, a Roth maximal count less than 12, and Roth counting time under 8 seconds offered high SP and +LR to rule in hypoxemia. Identifying patients with these features may be helpful in the remote assessment of COVID-19 outpatients. However, none of

these measures offered sufficient SN or –LR to help rule out hypoxemia– which is the more clinically important consideration for these patients.

Previous studies examining the correlation between subjective dyspnea and hypoxemia in other respiratory conditions have yielded inconsistent findings. The strongest confirmation of the potential diagnostic utility of dyspnea emerged from a study of 76 patients admitted to the emergency department with acute exacerbations of COPD, in which dyspnea scores exceeding 3 or 4 on a 5-stage scoring system were found to have a sensitivity of 93.5% for detecting hypoxemia.[13] Additionally, the mMRC Dyspnea Scale has been found to be significantly correlated with SpO2 in measurements obtained during exercise among patients with idiopathic pulmonary fibrosis.[18] Conversely, several other studies have shown no correlation between perceived dyspnea and hypoxemia in conditions such as advanced lung cancer, COPD, and palliative care patients.[14-15,19]

Discrepancies between respiratory rate and SpO2 in COVID-19 patients with acute respiratory failure have been highlighted previously, suggesting that a normal respiratory rate may belie profound hypoxemia in this setting.[20] High levels of anxiety may contribute to feelings of dyspnea in patients who are non-hypoxemic. There are also a growing number of case reports documenting silent hypoxemia among COVID-19 patients, where patients present with hypoxemia in the absence of respiratory symptoms.[9-10, 21] The underlying mechanism responsible for severe hypoxemia in the absence of dyspnea is not well elucidated. It has been postulated that this clinical picture may be consistent with a phenotype of COVID-19 pneumonia (L-phenotype) characterized by low elastance, low ventilation-perfusion ratio and near normal compliance.[11]

The relatively high compliance results in preserved gas volumes, while hypoxemia may result due to a ventilation-perfusion mismatch caused by impaired lung perfusion regulation and loss of hypoxic vasoconstriction.[22-23] Additionally, the absence of dyspnea despite severe hypoxemia may reflect pulmonary vaso-occlusive disease, whereby patients develop clinically silent microvascular thrombi in early stages of the disease, which if left untreated, results in worsening hypercoagulability and rapid clinical deterioration due to a thrombo-inflammatory cascade.[24-25] While at this point the exact mechanism remains speculative, our data suggest that the discrepancy between dyspnea and hypoxemia makes it difficult to accurately assess patients remotely and emphasizes the importance of SpO2 monitoring in order to avoid missing patients with developing respiratory failure.

This study has several limitations. The data collected for this study were from patients' initial pulse oximeter assessment and did not assess whether changes in dyspnea correlate with changes in SpO2 over time. This is a potentially important notion when monitoring patients who are (or are not) becoming increasingly dyspneic while self-isolating in their homes. While the number of patients included was sufficient for the primary analysis, they were insufficient for precise estimates of subgroups stratified by age, presence of lung disease, date of symptom onset, and for calculation of the diagnostic test characteristics at lower SpO2 cutoffs. Additionally, our study was limited to patients who were considered at high risk of severe disease, and it is possible that the diagnostic test characteristics might differ in younger and healthier patients. Lastly, pulse oximeters may have variable accuracy as individuals become increasingly hypoxic and are further impacted by individual patient characteristics; however, a perfect reference

standard of invasive blood oxygen measurement would be neither practical nor ethical in the outpatient setting.[26]

Our findings indicate that subjective dyspnea does not accurately capture hypoxemia in patients with COVID-19. Although some dyspnea scores have high specificity and positive likelihood ratios for identifying hypoxemia, none of these measures have sufficient sensitivity to rule out hypoxemia. Therefore, relying on surrogate measures of dyspnea alone is not sufficient to remotely monitor high-risk outpatients with COVID-19. Home SpO2 monitoring should be a mandatory component of remote management all high-risk outpatients with COVID-19.

References:

1. World Health Organization. Coronavirus disease 2019 (COVID-19) weekly update. https://www.who.int/publications/m/item/weekly-update-on-covid-19---23-october. Accessed October 24, 2020.
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020;323(13):1239–1242. Doi:10.1001/jama.2020.2648
3. Guan W, Ni Z, Yu Hu, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382(18):1708-1720. Doi:10.1056/NEJMoa2002032
4. Yan D, Wei L, Kui L, et al. Clinical characteristics of fatal and recovered cases of coronavirus disease 2019 (COVID-19) in Wuhan, China: a retrospective study. Chin. Med. J 2020;e-pub ahead of print. Doi:10.1097/CM9.0000000000000824
5. Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. BMJ 2020;368:m1091. Doi:10.1136/bmj.m1091
6. Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. BMJ 2020;368:m1182. Doi:https://doi.org/10.1136/bmj.m1182
7. Chorin E, Padegimas A, Havakuk O, et al. Assessment of respiratory distress by the Roth Score. Clin Cardiol 2019;39(11):636-639.
8. Friedman J, Calderón-Villarreal A, Bojorquez I, et al. Excess out-of-hospital mortality and declining oxygen saturation documented by EMS during the COVID-19 crisis in

Tijuana, Mexico. medRxiv 2020;e-pub ahead of print. Doi: https://doi.org/10.1101/2020.05.13.20098186.t

9. Ottestad W, Seim M, Maehlen JO. Covid-19 with silent hypoxemia. Tidsskr Nor Legeforen 2020;e-pub ahead of print. Doi:10.4045/tidsskr.20.0299
10. Wilkerson RG, Alder JD, Shah NG, Brown R. Silent hypoxia: A harbinger of clinical deterioration in patients with COVID-19. Am J Emerg Med 2020;e-pub ahead of print. Doi:https://doi.org/10.1016/j.ajem.2020.05.044
11. Gattinoni L, Chiumello D, Caironi P, et al. COVID-19 pneumonia: different respiratory treatments for different phenotypes? Intensive Care Med 2020; e-pub ahead of print. Doi: https://doi.org/10.1007/s00134-020-06033-2
12. Gondos T, Szabo V, Sarkany A, Sarkany A, Halasz G. Estimation of the severity of breathlessness in the emergency department: a dyspnea score. BMC Emergency Medicine 2016;17:13. Doi: 10.1186/s12873-017-0125-6
13. Guryay MS, Ceylan E, Gunay T, et al. Can Spirometry, pulse oximetry and dyspnea scoring reflect respiratory failure in patients with chronic obstructive pulmonary disease exacerbation? Medical Principles and Practice 2007;16:378-383. Doi:10.1159/000104812
14. Higashimoto Y, Honda N, Yamagata T. Exertional dyspnoea and cortical oxygenation in patients with COPD. European Respiratory Journal 2015;46:1615-1624. Doi:10.1183/13993003.00541-2015
15. Tanaka K, Akechi T, Okuyama T. Factors Correlated with Dyspnea in Advanced Lung Cancer Patients: Organic Causes and What Else? Journal of Pain and Symptom Management 2002;23(6):490-500. Doi:10.1016/s0885-3924(02)00400-1

16. Lam PW, Sehgal P, Andany N, et al. A virtual care program for outpatients diagnosed with COVID-19: A feasibility study. CMAJ Open 2020;8(2):E407-413.
17. Fletcher CM. Standardised questionnaire on respiratory symptoms: a statement prepared and approved by the MRC Committee on the Aetiology of Chronic Bronchitis (MRC breathlessness score). BMJ 1960;2:1665.
18. Manali ED, Lyberopoulos P, Triantafillidou C, et al. MRC chronic dyspnea scale: relationships with cardiopulmonary exercise testing and 6-minute walk test in idiopathic pulmonary fibrosis patients: a prospective study. BMC Pulmonary Medicine 2010;10:32.
19. Clemens KE, Quednau I, Klaschik E. Use of oxygen and opioids in the palliation of dyspnoea in hypoxic and non-hypoxic palliative care patients: a prospective study. Support Care Cancer 2009;17:367-377.
20. Jouffroy R, Jost D, Prunet B. Prehospital pulse oximetry: a red flag for early detection of silent hypoxemia in COVID-19 patients. Critical Care 2020;24:313.
21. Tobin MJ, Laghi F, Jubran A. "Why COVID-19 silent hypoxemia is baffling to physicians. Annals of the American Thoracic Society 2020; e-pub ahead of print. doi:10.1164/rccm.202006-2157CP
22. Gattinoni L, Coppola S, Cressoni M, et al. COVID-19 does not lead to a "typical" acute respiratory distress syndrome. Am J Respir Crit Care Med 2020;201(10):1299-1300.
23. Gattinoni L, Chiumello D, Rossi S. COVID-19 pneumonia: ARDS or not? Critical Care 2020;24:154
24. Low T, Cherian R, Lim SL, et al. Rethinking COVID-19 'pneumonia'-is this primarily a vaso-occlusive disease, and can early anticoagulation save the ventilator famine. Pulmonary Circulation 2020;10(3):1-3.

25. Couzel-Frankel J. The mystery of the pandemic's 'happy hypoxia'. Science 2020;368(6490);455-456.

26. Luks AM, Swenson ER. Pulse oximetry for monitoring patients with COVID-19 at home: potential pitfalls and practical guidance. Annals of the American Thoracic Society 2020; e-pub ahead of print. doi: https://doi.org/10.1513/AnnalsATS.202005-418FR

Figure Legends:

Figure 1. Comparison of SpO2 and measures of subjective dyspnea. A) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients who reported dyspnea and those who did not. B) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients with various mMRC Dyspnea Scale scores. The width of each plot is proportional to the number of patients with the respective SpO2 (represented by black dots). The median SpO2 is indicated by the central horizontal black line and the dotted lines correspond to the interquartile range.

Table 1: Demographics and Clinical Characteristics Among Outpatients with COVID-19 Monitored with Home Oxygen Saturation Devices	
Demographic information	**No. (%)**
Total No.	89
Age, median (IQR), y	52 (38-64)
Sex	
Female	57(64)
Male	32 (36)
Pregnant	6 (7)
Days from symptom onset to clinical assessment, median (IQR)	6 (3-8)
Comorbidities	
Cardiac disease	7 (8)
Chronic lung disease	3 (3)
Asthma	14 (16)
Chronic kidney disease	7 (8)
Moderate/severe liver disease	2 (2)
Chronic neurological issues	5 (6)
Malignancy	14 (16)
Chronic hematological disease	7 (8)
Diabetes	15 (17)
Hypertension	32 (36)
Rheumatic disorder	3 (3)
Malnutrition	1 (1)
Obesity	18 (20)
None	29 (33)
Signs and symptoms on intake assessment	
Fever	35 (39)
Sore Throat	28 (31)
Runny Nose	29 (33)
Cough	56 (63)
Shortness of Breath	23 (26)
Chills/Rigors	30 (34)
Conjunctivitis	10 (11)
Ear Pain	7 (8)
Anosmia	21 (24)
Dysgeusia	25 (28)
Sputum	10 (11)
Hemoptysis	0
Wheezing	7 (8)
Chest Pain	19 (21)
Myalgia	40 (45)
Arthralgia	16 (18)
Abdominal Pain	14 (16)

Nausea/Vomiting	22 (25)
Diarrhea	25 (28)
Adenopathy	0
Rash	1 (1)
Fatigue/Malaise	59 (66)
Headache	37 (42)
Confusion	5 (6)
Depression/Anxiety	15 (17)
Insomnia	19 (21)
Anorexia	33 (37)
Laboratory findings at admission, median (IQR)	
Leukocytes, $x10^9$/L (n=21)	5.9 (4.2-7.5)
Lymphocytes, $x10^9$/L (n=21)	1.1 (0.6-1.3)
Lactate Dehydrogenase, IU/L (n=4)	216.0 (63.4-277.0)
D-dimer, mcg/L (n=5)	906.0 (555.0-1082.5)
High-sensitivity Troponin T ng/L (n=11)	9.7 (6.0-10.0)
Ferritin, mcg/L (n=3)	1644.0 (153.5-2082.5)
Chest radiography done	29 (33)
Abnormal	23 (26)
Bilateral infiltrates	18 (20)
Outcome	
ICU admission	5 (6)
Length of ICU stay, median (IQR), days	6 (2-11)
Intubation	1 (1)
Duration of intubation, days	15 (17)
Hospitalization	11 (12)
Duration of hospitalization, median (IQR), days	3 (2.5-7)
Multiple hospitalizations	2 (2)
Death (within 30 days of diagnosis)	0

Table 2: Association of Dyspnea Measurements with Detection of Hypoxemia			
Dyspnea measurement	**Non-hypoxic patients (O2 sat ≥95%)**	**Hypoxic Patients (O2 sat <95%)**	**p-value**
Shortness of breath	14	8	0.046*
Breathing faster at rest	3	0	1.00
Breathing harder than normal	2	0	1.00
More breathless today than yesterday	3	1	0.57
mMRC Dyspnea scale:			
>0	9	7	0.014*
>1	1	5	0.001*
>2	1	5	0.001*
>3	1	1	0.37
Roth Score: Maximum Count			
<12	0	1	0.21
<15	3	2	0.27
<20	5	2	0.60
<28	6	2	1.00
Roth Score: Count time			
<8 sec	1	1	0.39
<15 sec	10	2	0.60
<20 sec	12	2	0.27
<25 sec	13	3	0.53

Table 3: Diagnostic Accuracy of Dyspnea Measurements in the Detection of Hypoxemia: Estimation of test characteristics including sensitivity (SN), specificity, negative (NPV) and positive (PPV) predictive value and negative (-LR) and positive (+LR) likelihood ratio

Dyspnea measurement	SpO2 Cutoff	SN % (95%CI)	SP % (95%CI)	NPV % (95%CI)	PPV % (95%CI)	-LR (95%CI)	+LR (95%CI)
Shortness of breath	≤97%	25 (16-37)	75 (47-92)	19 (10-30)	82 (59-94)	1.0 (0.7-1.4)	1.0 (0.4-2.6)
	≤96%	27 (17-41)	78 (60-90)	39 (27-51)	68 (45-85)	0.9 (0.7-1.2)	1.3 (0.6-2.7)
	≤95%	40 (23-59)	83 (70-91)	72 (60-82)	55 (33-75)	0.7 (0.5-1.0)	2.3 (1.1-4.7)
	≤94%	47 (24-72)	80 (68-88)	86 (75-93)	36 (18-59)	0.7 (0.4-1.1)	2.4 (1.2-4.7)
	≤93%	46 (18-75)	78 (66-86)	91 (80-96)	23 (9-46)	0.7 (0.4-1.2)	2.0 (0.9-4.4)
	≤92%	50 (19-81)	77 (66-85)	95 (86-99)	14 (4-36)	0.6 (0.3-1.5)	2.1 (0.9-5.2)
Breathing faster at rest	≤94%	0 (0-24)	96 (87-99)	81 (70-88)	0 (0-69)	1.0 (1.0-1.1)	0
Breathing harder than normal	≤94%	0 (0-24)	97 (89-100)	81 (71-88)	0 (0-80)	1.0 (1.0-1.1)	0
More breathless today than yesterday	≤94%	6.2 (0.3-32)	96 (87-99)	82 (71-89)	25 (1-78)	1.0 (0.9-1.1)	1.5 (0.2-13.1)
mMRC Dyspnea scale:							
>0	≤94%	54 (26-80)	82 (68-91)	87 (74-95)	44 (21-69)	0.6 (0.3-1.0)	3.0 (1.4-6.5)
>1	≤94%	39 (15-68)	98 (88-100)	86 (74-93)	83 (37-99)	0.6 (0.4-1.0)	19.2 (2.4-151)
>2	≤94%	39 (15-68)	98 (88-100)	86 (74-93)	83 (37-99)	0.6 (0.4-1.0)	19.2 (2.4-151)
>3	≤94%	8 (0.4-38)	98 (88-100)	80 (68-89)	50 (9.5-91)	0.9 (0.8-1.1)	3.9 (0.3-57.4)
Roth Score: Maximum Count							
<12	≤94%	25 (1.3-78)	100 (75-100)	83 (58-96)	100 (6-100)	0.8 (0.4-1.3)	N/A
<15	≤94%	50 (15-85)	80 (51-95)	86 (56-98)	40 (7-83)	0.6 (0.2-1.7)	2.5 (0.6-10.2)
<20	≤94%	50 (15-85)	67 (39-87)	83 (51-97)	29 (5.1-70)	0.8 (0.3-2.1)	1.5 (0.5-5.1)
<28	≤94%	7.7 (0.4-38)	94 (82-98)	79 (66-88)	25 (1.3-78)	1.0 (0.3-2.4)	1.3 (0.4-4.0)
Roth Score: Count time							
<8 sec	≤94%	25 (1.3-78)	93 (66-100)	82 (56-95)	50 (10-91)	0.8 (0.5-1.4)	3.8 (0.3-47.7)
<15 sec	≤94%	50 (15-85)	33 (13-61)	71 (30-95)	17 (2.9-49)	1.5 (0.5-5.1)	0.8 (0.3-2.1)
<20 sec	≤94%	50 (15-85)	20 (5.3-49)	60 (17-93)	14 (2.5-44)	2.5 (0.6-10.2)	0.6 (0.2-1.7)
<25 sec	≤94%	75 (22-99)	13 (2.3-42)	67 (13-98)	19 (5.0-46)	1.9 (0.2-15.8)	0.9 (0.5-1.6)

Table 4: Diagnostic accuracy of the presence of dyspnea in the detection of hypoxemia stratified based on patient characteristics, including age, presence vs absence of underlying lung disease, and date from symptom onset

	SN % (95%CI)	SP % (95%CI)	NPV % (95%CI)	PPV % (95%CI)	-LR (95%CI)	+LR (95%CI)
Age						
<60 y	75 (36-96)	70 (54-83)	94 (78-99)	32 (14-57)	0.4 (0.1-1.2)	2.5 (1.3-4.5)
≥60 y	22 (3.9-60)	96 (79-100)	79 (61-90)	67 (13-98)	0.8 (0.6-1.2)	6.0 (0.6-58.6)
Underlying lung disease						
Yes	100 (20-100)	80 (51-95)	100 (70-100)	40 (7.3-83)	0	5.0 (1.8-13.8)
No	40 (18-67)	80 (67-89)	83 (70-92)	35 (15-61)	0.8 (0.5-1.2)	2.0 (0.9-4.5)
Days from symptom onset						
<7 days	50 (19-81)	83 (68-93)	92 (78-98)	30 (8.1-65)	0.6 (0.3-1.4)	3.0 (1.1-8.6)
≥7 days	50 (24-76)	71 (48-88)	75 (51-90)	46 (18-75)	0.7 (0.4-1.4)	1.8 (0.7-4.4)

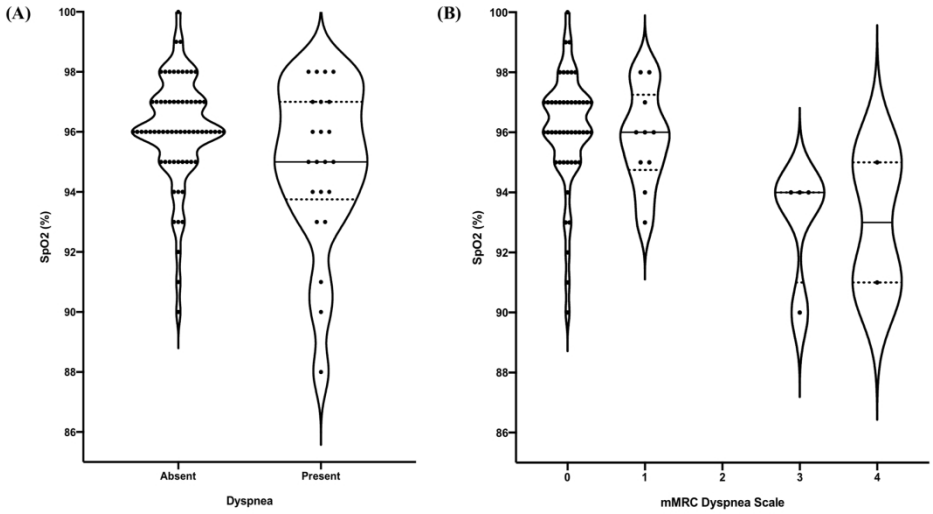

Figure 1. Comparison of SpO2 and measures of subjective dyspnea. A) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients who reported dyspnea and those who did not. B) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients with various mMRC Dyspnea Scale scores. The width of each plot is proportional to the number of patients with the respective SpO2 (represented by black dots). The median SpO2 is indicated by the central horizontal black line and the dotted lines correspond to the interquartile range.

BMJ Open

The Diagnostic Accuracy of Subjective Dyspnea in Detecting Hypoxemia Among Outpatients with COVID-19: a retrospective cohort study

Journal:	*BMJ Open*
Manuscript ID	bmjopen-2020-046282.R1
Article Type:	Original research
Date Submitted by the Author:	23-Jan-2021
Complete List of Authors:	Berezin, Linor; University of Toronto Faculty of Medicine Zhabokritsky, Alice; University of Toronto Faculty of Medicine Andany, Nisha; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto, Faculty of Medicine Chan, Adrienne ; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Estrada-Codecido, Jose; Sunnybrook Health Sciences Centre, Division of Infectious Diseases Gershon, Andrea; Sunnybrook Health Sciences Centre, Division of Respirology; University of Toronto Faculty of Medicine Lam, Philip; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Leis, Jerome; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine MacPhee, Scott; Sunnybrook Health Sciences Centre, Department of Nursing Mubareka, Samira; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto, Department of Laboratory Medicine and Pathology Simor, Andrew; Sunnybrook Health Sciences Centre, Division of Infectious Diseases; University of Toronto Faculty of Medicine Daneman, Nick; Sunnybrook Health Sciences Centre; University of Toronto Faculty of Medicine
Primary Subject Heading:	Infectious diseases
Secondary Subject Heading:	Medical management
Keywords:	COVID-19, INFECTIOUS DISEASES, Public health < INFECTIOUS DISEASES

I, the Submitting Author has the right to grant and does grant on behalf of all authors of the Work (as defined in the below author licence), an exclusive licence and/or a non-exclusive licence for contributions from authors who are: i) UK Crown employees; ii) where BMJ has agreed a CC-BY licence shall apply, and/or iii) in accordance with the terms applicable for US Federal Government officers or employees acting as part of their official duties; on a worldwide, perpetual, irrevocable, royalty-free basis to BMJ Publishing Group Ltd ("BMJ") its licensees and where the relevant Journal is co-owned by BMJ to the co-owners of the Journal, to publish the Work in this journal and any other BMJ products and to exploit all rights, as set out in our licence.

The Submitting Author accepts and understands that any supply made under these terms is made by BMJ to the Submitting Author unless you are acting as an employee on behalf of your employer or a postgraduate student of an affiliated institution which is paying any applicable article publishing charge ("APC") for Open Access articles. Where the Submitting Author wishes to make the Work available on an Open Access basis (and intends to pay the relevant APC), the terms of reuse of such Open Access shall be governed by a Creative Commons licence – details of these licences and which Creative Commons licence will apply to this Work are set out in our licence referred to above.

Other than as permitted in any relevant BMJ Author's Self Archiving Policies, I confirm this Work has not been accepted for publication elsewhere, is not being considered for publication elsewhere and does not duplicate material already published. I confirm all authors consent to publication of this Work and authorise the granting of this licence.

The Diagnostic Accuracy of Subjective Dyspnea in Detecting Hypoxemia Among Outpatients with COVID-19: a retrospective cohort study

Author list:

Linor Berezin[a], Alice Zhabokritsky[a], Nisha Andany[a,b], Adrienne K. Chan[a,b], Jose Estrada-Codecido[b], Andrea Gershon[a,c], Philip W. Lam[a,b], Jerome A. Leis[a,b], Scott MacPhee[d], Samira Mubareka[b,e], Andrew E. Simor[a,b], Nick Daneman[a,b].

Author Affiliations:

[a] Faculty of Medicine, University of Toronto, Toronto, Canada

[b] Division of Infectious Diseases, Sunnybrook Health Sciences Centre, University of Toronto

[c] Division of Respirology, Sunnybrook Health Sciences Centre, University of Toronto

[d] Department of Nursing, Sunnybrook Health Sciences Centre, University of Toronto

[e] Department of Laboratory Medicine and Pathology, University of Toronto, Toronto, Canada

Corresponding author:

Nick Daneman, MD, FRCPC, MSc

Division of Infectious Diseases, Sunnybrook Health Sciences Centre, University of Toronto,

2075 Bayview Avenue, Toronto, ON, M4N 2M5, Canada

nick.daneman@sunnybrook.ca

Key Words: COVID-19, Dyspnea, hypoxemia, oxygen saturation, monitoring, virtual care

Word Count: 2831

Author Contributions: The authors all stand behind the conclusions of this manuscript, agree to be accountable for all aspects of the work, and support its publication. LB contributed to the planning, conception, and study design, data analysis, interpretation of the data, and reporting of the work. AZ contributed to data analysis, interpretation of the data, and reporting of the work. NA contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. AKC contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. JE contributed to data analysis, interpretation of the data, and reporting of the work. AG contributed to the interpretation of the data and reporting of the work. PWL contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. JAL contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. SM contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. SM contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. AES contributed to the planning, conception, study design, study conduct, data acquisition, and reporting of the work. ND contributed to the planning, conception, and study design, study conduct, data acquisition, data analysis, interpretation of the data, and reporting of the work. All authors contributed to the manuscript preparation and have given approval for its submission.

Conflicts of Interest: The authors report no financial or other conflicts of interest.

Abstract:

Objectives: The majority of patients with mild-to-moderate COVID-19 can be managed using virtual care. Dyspnea is challenging to assess remotely, and the accuracy of subjective dyspnea measures in capturing hypoxemia have not been formally evaluated for COVID-19. We explored the accuracy of subjective dyspnea in diagnosing hypoxemia in COVID-19 patients.

Methods: This is a retrospective cohort study of consecutive outpatients with COVID-19 who met criteria for home oxygen saturation monitoring at a university-affiliated acute care hospital in Toronto, Canada from April 3, 2020 to September 13, 2020. Dyspnea measures were treated as diagnostic tests, and we determined their sensitivity (SN), specificity (SP), negative/positive predictive value (NPV/PPV), and positive/negative likelihood ratios (+LR/-LR) for detecting hypoxemia. In the primary analysis, hypoxemia was defined by oxygen saturation <95%; the diagnostic accuracy of subjective dyspnea was also assessed across a range of oxygen saturation cutoffs from 92% to 97%.

Results: During the study period 89/501 (17.8%) of patients met criteria for home oxygen saturation monitoring, and of these 17/89 (19.1%) were diagnosed with hypoxemia. The presence/absence of dyspnea had limited accuracy for diagnosing hypoxemia, with SN 47% (95%CI 24-72%), SP 80% (68%-88%), NPV 86% (75%-93%), PPV 36% (18%-59%), +LR 2.4 (1.2-4.7), -LR 0.7 (0.4-1.1). The SN of dyspnea was 50% (95%CI 19-81) when a cutoff of ≤92% was used to define hypoxemia. A modified Medical Research Council (mMRC) dyspnea score >1 (SP 98%, 95%CI 88%-100%), Roth Maximal Count <12 (SP 100%, 95%CI 75-100%), and Roth Counting time < 8 seconds (SP 93%, 95%CI 66%-100%) had high SP that could be used to rule in hypoxemia, but displayed low SN (≤50%).

Conclusions: Subjective dyspnea measures have inadequate accuracy for ruling out hypoxemia in high-risk patients with COVID-19. Safe home management of patients with COVID-19 should incorporate home oxygenation saturation monitoring.

Strengths and limitations of the study:

- This is the first study to evaluate the diagnostic accuracy of subjective dyspnea in detecting hypoxemia in the setting of COVID-19.
- The diagnostic accuracy of patient-reported presence of dyspnea in capturing hypoxemia was evaluated across a range of SpO2 cutoffs from 92 to 97% and stratified based on age, presence of lung disease, and date of symptom onset.
- Subgroup analysis of the diagnostic accuracy of dyspnea as measured by objectives measures such as the mMRC dyspnea scale Roth Score is included.
- Methodological limitations of the study include the retrospective study design and small sample size.
- This study was limited to patients who were considered high risk for severe COVID-19, and the data collected for this study were from single patient assessments and did not assess whether changes in dyspnea correlate with changes in SpO2 over time.

Introduction:

As of January 19, 2021, there have been more than 93 million laboratory-confirmed novel coronavirus disease 2019 (COVID-19) cases and 2 million deaths documented globally.[1] The spectrum of disease of COVID-19 ranges from asymptomatic or mild symptoms, to severe respiratory failure and death.[2] Approximately 20% of patients with COVID-19 experience dyspnea, which is more commonly associated with severe disease.[3] Fatal cases of COVID-19 have higher rates of dyspnea, lower blood oxygen saturation (SpO2), and greater rates of complications such as acute respiratory distress syndrome.[4,5]

In an effort to reduce avoidable hospitalizations, health care contacts, and transmission, most patients with COVID-19 can be managed in the community using virtual healthcare platforms, and transferred to hospital only if they develop progressive respiratory disease.[6] Subjective dyspnea can be assessed remotely using patient interview, and augmented by surrogate measures such as the Roth Score[7] and modified Medical Research Council (mMRC) Dyspnea Scale.[8] However, the accuracy of these measures has not been formally evaluated in the context of COVID-19.[6,7] Of great concern is the risk of false reassurance if patients develop hypoxemia without the subjective sensation of dyspnea. “Silent hypoxemia”, or low SpO2 in the absence of dyspnea, has been reported in the setting of COVID-19 and clinicians have speculated that it may be associated with increased out-of-hospital mortality;[9] case reports have described patients presenting to hospital with rapid deterioration and respiratory failure without signs of respiratory distress.[10-12]

Previous studies of the utility of dyspnea measurement in diagnosing hypoxemia in other respiratory conditions such as chronic obstructive pulmonary disease (COPD), congestive heart failure and lung cancer have yielded conflicting results;[13-16] The association has not been studied during the COVID-19 pandemic despite the highly publicized concept of "silent hypoxemia". Therefore, we sought to determine the diagnostic accuracy of subjective dyspnea measures in diagnosing hypoxemia among a cohort of outpatients with COVID-19.

Methods:

Study Participants

All consecutive patients with laboratory-confirmed COVID-19 followed as outpatients by the Sunnybrook Health Sciences Centre COVID-19 Expansion to Outpatients (COVIDEO) virtual care service from April 3, 2020 to September 13, 2020 were included in this retrospective cohort study. The patients were diagnosed based on a positive mid-turbinate or nasopharyngeal swab for COVID-19 RNA detected by real-time polymerase chain reaction. COVIDEO is a virtual care model for monitoring of outpatients with COVID-19 at Sunnybrook Health Sciences Centre, and is the basis of similar programs at other hospitals.[17] Patients were contacted by an infectious diseases physician for assessment and monitoring either by telephone or through the Ontario Telemedicine Network virtual platform.

A portable pulse oximeter was delivered to the homes of high-risk patients as defined by age > 60 years, pregnancy, extensive comorbidities, or presence of cardio-respiratory symptoms, such as chest pain or dyspnea. This study was approved by the institutional review board at

Sunnybrook Health Sciences as minimal risk research, using data collected for routine clinical practice, and the requirement for informed consent was waived.

Patient and Public Involvement

No patient involved.

Data Collection

The demographic characteristics, clinical data, measures of subjective dyspnea (presence of shortness of breath, mMRC dyspnea scale score, Roth Score), physical exam findings, and SpO2 readings for study participants were collected from electronic medical records by one investigator (either A.Z or S.Ma.). For the analysis, values were obtained from the patient's initial virtual care assessment with a pulse oximeter, and subjective dyspnea measures were taken at the same time as the objective measure of hypoxia.

Predictor Variables

The primary predictor of interest was patient-reported presence of dyspnea. Secondary predictor variables were patient-reported breathing faster at rest, breathing harder than normal, feeling more breathless today than yesterday, as well as dyspnea as measured by the mMRC Dyspnea Scale and the Roth Score.

The mMRC Dyspnea Scale has been studied extensively in a variety of respiratory conditions.[8] It is comprised of five categories that describe the degree of activity limitation due to worsening breathlessness. The participant assigns themselves a score ranging from 0 to 4 based on their

perception of which activities result in dyspnea, with higher scores indicating a greater impairment in their ability to perform daily activities.

The Roth Score is a tool for quantifying the severity of dyspnea, in which the patient is asked to count audibly to 30 in their native language, and the maximal count and counting time are recorded. A prior validation study demonstrated a strong positive correlation between pulse oximetry measurement and both counting time ($r = 0.59$; $P < 0.001$) and maximal count ($r = 0.67$; $P < 0.001$) achieved in one breath.[7]

Outcomes

The reference measure was SpO2 as measured by a ChoiceMMed pulse oximeter (model MD300C20). In the primary outcome definition, hypoxemia was considered to be present if SpO2 was < 95% in order to provide sufficient power to estimate the diagnostic test characteristics. Secondary outcomes included hypoxemia cut-offs varying from 92 to 97%. Patients received instructions on correct pulse oximeter use and were told to wait 5 to 10 seconds for readings to calibrate prior to recording the SpO2 measurements.

Statistical Analysis

In the primary analysis, the subjective dyspnea measures were treated as diagnostic tests, and the specificity (SP), sensitivity (SN), positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (+LR) and negative likelihood ratio (-LR) were determined in order to evaluate the predictive value in detecting hypoxemia. For the continuous predictors, the test characteristics were provided across a range of different score thresholds.

Diagnostic test characteristics of the primary dyspnea measure were also determined in subgroups stratified based on patient characteristics, including (1) age <60 vs >60 years, (2) presence vs absence of underlying lung disease, and date from symptom onset (<7 vs ≥7 days). The Wilson method with continuity correction was used to calculate 95% confidence intervals in order to avoid a negative lower limit.

A secondary analysis examined the strength of association between the presence of dyspnea and hypoxemia with a χ^2 test or Fisher exact test (for sample sizes <5) with dyspnea treated as a binary variable (present or absent). This relationship is represented by a violin plot. In additional analyses, a correlation coefficient was calculated to assess whether there was an association between the participants' Roth Scores and their oxygen saturation measurements. These associations were displayed graphically with a scatter plot (Figure S1 and S2). The relationship between the mMRC Dyspnea Scale and hypoxemia was analyzed with a χ^2 test and represented by violin plot.

All analyses were conducted in SAS Statistical Software V.9.3 (Cary, North Carolina, USA). For all statistical analyses, $P < 0.05$ was considered significant.

Sample Size Calculation

The primary test characteristic of interest was the SN of dyspnea as a test for hypoxemia. The sample size was estimated based on a test of single proportion, namely SN. It was estimated that

62 patients would be required in order to estimate a SN with a 10% margin of error and 95% confidence if the true sensitivity was 80%.

Results:

Demographic and Clinical Characteristics of Outpatients with COVID-19

From April 3 to Sept 13, 2020, a total of 501 patients with COVID-19 were followed by COVIDEO. Of these patients, 89 (17.8%) met criteria for home oxygen saturation monitoring (age > 60 years, pregnancy, extensive comorbidities, or presence of cardio-respiratory symptoms). One patient was lost to follow up after provision of the oxygen monitoring device.

Overall, the median age of patients was 52 years (interquartile range (IQR) 38-64 years) and 57 patients (64%) were female (Table 1). The median number of days from symptom onset to clinical assessment was 6 (IQR 3-8). Among these patients, the most common comorbidities were hypertension (36%), obesity (20%), diabetes (17%), asthma (16%), and malignancy (16%). Twenty-nine patients (33%) had no comorbidities. The most common symptoms reported on intake assessment were fatigue/malaise (66%), cough (63%), and myalgias (45%). While the patients were being followed by COVIDEO, 11 (12%) required hospitalization, with a median duration of hospitalization of 3 days (IQR 2.5-7). Five (6%) patients were admitted to the intensive care unit (ICU), and the median length of ICU stay was 6 days (IQR 2-11). One patient was intubated, and no patients died within 30 days of their COVID-19 diagnosis.

Association of Dyspnea Measurements with Detection of Hypoxemia

A total of 17 (19.1%) patients were diagnosed with hypoxemia. Hypoxemia was significantly associated with the presence of dyspnea (p= 0.046), mMRC Dyspnea Scale score over 0 (p= 0.014), over 1 (p= 0.001) and over 2 (p= 0.001) (Table 2). Weak associations were identified between patients' Roth Scores and their oxygen saturation measurements for maximum count (r =0.29; p=0.23) and counting time (r =0.12; p=0.617), respectively. The distribution of SpO2 (%) values in COVID-19 outpatients who reported dyspnea and with various mMRC Dyspnea Scale scores is shown in Figure 1.

Diagnostic Accuracy of Dyspnea Measurements in the Detection of Hypoxemia:

The presence or absence of subjective dyspnea had a SN 47% (95%CI 24-72%), SP 80% (68%-88%), NPV 86% (75%-93%), PPV 36% (18%-59%), +LR 2.4 (1.2-4.7), -LR 0.7 (0.4-1.1) for diagnosing hypoxemia (Table 3). The presence of subjective dyspnea had lower SN (25% [16%-37%]), 27% [17%-41%], 40% (23%-59%) for detecting hypoxemia as defined by thresholds of ≤97%, ≤96%, and ≤95%, respectively. At a lower SpO2 threshold of ≤92%, the SN increased only slightly to 50% (19%-81%). The other binary measures of subjective dyspnea, including breathing faster at rest, breathing harder than normal, and feeling more breathless than the day before had lower SN (0% [0%-24%], 0% [0%-24%], and 6.2% [0.3%-32%], respectively), and higher SP (96% [87%-99%], 97% [89%-100%], and 96% [87%-99%], respectively) (Table 3).

mMRC Dyspnea Scale scores were recorded and available for 63 patients (70.8%). An mMRC Dyspnea Scale score of greater than 0 was determined to have a SN of 54% (26%-80%), SP 82% (68%-91%), NPV 87% (74%-95%), PPV 44% (21-69%), -LR 0.6 (0.3-1.0), and +LR 3.0 (1.4-6.5) for the detection of hypoxemia. At higher cutoff values, the SN of the mMRC Dyspnea

Scale was reduced to 39% (15%-68%) for scores greater than 1 and 2 and 8% (0.4%-38%) for scores greater than 3. The SP for mMRC Dyspnea Scale scores greater than 1, 2, and 3 at capturing hypoxemia was 98% (88%-100%).

Roth Scores were available for 19 patients (29.7%). The Roth Score had a higher SN at higher cutoff values for counting time. A maximum count of less than 12 had a SN of 25% (1.3%-78%), SP 100% (75%-100%), NPV 83% (58%-96%), PPV 100% (6%-100%), and a -LR of 0.75 (0.4-1.3).

The diagnostic test with the highest SN for diagnosing hypoxemia was a Roth score maximum counting time of less than 25 seconds, which still had a SN of only 75% (22%-99%), and inadequate SP 13% (2.3%-42%), NPV 67% (13%-98%), PPV 19% (5.0%-46%), -LR 1.88 (0.2-16), and +LR 0.87 (0.48-1.6). When all subjective dyspnea predictors are combined in a single variable, the SN is 59% (34%-81%), SP 67% (55%-78%), NPV 87% (75%-94%), PPV 30% (16%-49%), -LR 0.6 (0.3-1.1), and +LR 1.8 (1.0-3.0).

The diagnostic accuracy of dyspnea presence in the detection of hypoxemia was most impacted when stratified by the presence of underlying lung disease. In the patients with underlying lung disease, the SN and SP of the presence of dyspnea in detecting hypoxemia was 100% (20%-100%) and 80% (51%-95%), respectively. A lower SN (22% [3.9%-60%]) and high SP (96% [79%-100%]) was observed for patients over 60 years when results were stratified based on age. Stratifying based on days from symptom onset did not impact the diagnostic accuracy of dyspnea in detecting hypoxemia (Table 4).

Discussion:

To our knowledge, this is the first study to evaluate the diagnostic accuracy of subjective dyspnea in detecting hypoxemia in the setting of COVID-19. Self-reported shortness of breath has very limited utility for detecting hypoxemia, with a SN of only 47% and SP of only 80% for detecting SpO2 levels below 95%. Using a lower SpO2 threshold of less than 93% did not meaningfully improve the SN of subjective dyspnea in diagnosing hypoxemia (SN 50%). Other binary measures of subjective dyspnea, including breathing faster at rest, breathing harder than normal, and feeling more breathless than yesterday offered high SP. Similarly, an mMRC Dyspnea Score exceeding 1, a Roth maximal count less than 12, and Roth counting time under 8 seconds offered high SP and +LR to rule in hypoxemia. Identifying patients with these features may be helpful in the remote assessment of COVID-19 outpatients. However, none of these measures offered sufficient SN or –LR to help rule out hypoxemia– which is the more clinically important consideration for these patients. When stratified based on the presence of underlying lung disease, the SN and SP of the presence of subjective dyspnea in detecting hypoxemia increased to 100% and 80%, respectively, suggesting that underlying lung disease may be a useful clinical feature for ruling out hypoxemia in COVID-19 patients on the basis of subjective reports of dyspnea. Even when all variables were combined into a single maximally sensitive predictor, the SN was just 59%.

Previous studies examining the correlation between subjective dyspnea and hypoxemia in other respiratory conditions have yielded inconsistent findings. The strongest confirmation of the potential diagnostic utility of dyspnea emerged from a study of 76 patients admitted to the emergency department with acute exacerbations of COPD, in which dyspnea scores exceeding 3 or 4 on a 5-stage scoring system were found to have a sensitivity of 93.5% for detecting hypoxemia.[14] Additionally, the mMRC Dyspnea Scale has been found to be significantly correlated with SpO2 in measurements obtained during exercise among patients with idiopathic pulmonary fibrosis.[18] Conversely, several other studies have shown no correlation between perceived dyspnea and hypoxemia in conditions such as advanced lung cancer, COPD, and palliative care patients.[15-16,19] While previous studies show variable relationships between dyspnea and hypoxemia in various respiratory pathologies, our study shows that neither binary measures of subjective dyspnea, the mMRC Dyspnea Scale, or the Roth Score can be used to diagnose hypoxemia in the setting of COVID-19.

Discrepancies between respiratory rate and SpO2 in COVID-19 patients with acute respiratory failure have been highlighted previously, suggesting that a normal respiratory rate may belie profound hypoxemia in this setting.[20] High levels of anxiety may contribute to feelings of dyspnea in patients who are non-hypoxemic. There are also a growing number of case reports documenting silent hypoxemia among COVID-19 patients, where patients present with hypoxemia in the absence of respiratory symptoms.[10-11, 21] The underlying mechanism responsible for severe hypoxemia in the absence of dyspnea is not well elucidated. It has been postulated that this clinical picture may be consistent with a phenotype of COVID-19 pneumonia (L-phenotype) characterized by low elastance, low ventilation-perfusion ratio and near normal

compliance.[12] The relatively high compliance results in preserved gas volumes, while hypoxemia may result due to a ventilation-perfusion mismatch caused by impaired lung perfusion regulation and loss of hypoxic vasoconstriction. [22-23] Additionally, the absence of dyspnea despite severe hypoxemia may reflect pulmonary vaso-occlusive disease, whereby patients develop clinically silent microvascular thrombi in early stages of the disease, which if left untreated, results in worsening hypercoagulability and rapid clinical deterioration due to a thrombo-inflammatory cascade.[24-25] While at this point the exact mechanism remains speculative, our data suggest that the discrepancy between dyspnea and hypoxemia makes it difficult to accurately assess patients remotely and emphasizes the importance of SpO2 monitoring in order to avoid missing patients with developing respiratory failure.

This study has several limitations. The data collected for this study were from patients' initial pulse oximeter assessment and did not assess whether changes in dyspnea correlate with changes in SpO2 over time. This is a potentially important notion when monitoring patients who are (or are not) becoming increasingly dyspneic while self-isolating in their homes. While the number of patients included was sufficient for the primary analysis, they were insufficient for precise estimates of subgroups stratified by age, presence of lung disease, date of symptom onset, and for calculation of the diagnostic test characteristics at lower SpO2 cutoffs. In our study, less than 20% of included patients were diagnosed with hypoxemia. While this represents a small sample of patients with hypoxemia, it is clear that in order to prevent missing any patients with hypoxemia who require admission all high risk patients require oxygen saturation monitoring. Additionally, our study was limited to patients who were considered at high risk of severe disease, and it is possible that the diagnostic test characteristics might differ in younger and

healthier patients. Lastly, pulse oximeters may have variable accuracy as individuals become increasingly hypoxic and are further impacted by individual patient characteristics; however, a perfect reference standard of invasive blood oxygen measurement would be neither practical nor ethical in the outpatient setting.[26]

Conclusions:

Our findings indicate that subjective dyspnea does not accurately capture hypoxemia in patients with COVID-19. Although some dyspnea scores have high specificity and positive likelihood ratios for identifying hypoxemia, none of these measures have sufficient sensitivity to rule out hypoxemia. Therefore, relying on surrogate measures of dyspnea alone is not sufficient to remotely monitor high-risk outpatients with COVID-19. Home SpO2 monitoring should be a mandatory component of remote management all high-risk outpatients with COVID-19.

Competing Interests:

The authors have no competing interests to declare.

Funding:

There is no funding to report fo this submission.

Data Sharing Statement:

All data relevant to the study are included in the article or uploaded as supprelmentary information. Data are available upon reasonable request. Deidentified data are available upon reasonable request.

References:

1. World Health Organization. Coronavirus disease 2019 (COVID-19) weekly epidemiological update. https://www.who.int/publications/m/item/weekly-epidemiological-update---19-january-2021. Accessed January 21, 2021.
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. JAMA 2020;323(13):1239–1242. Doi:10.1001/jama.2020.2648
3. Guan W, Ni Z, Yu Hu, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med 2020;382(18):1708-1720. Doi:10.1056/NEJMoa2002032
4. Yan D, Wei L, Kui L, et al. Clinical characteristics of fatal and recovered cases of coronavirus disease 2019 (COVID-19) in Wuhan, China: a retrospective study. Chin. Med. J 2020;e-pub ahead of print. Doi:10.1097/CM9.0000000000000824
5. Chen T, Wu D, Chen H, et al. Clinical characteristics of 113 deceased patients with coronavirus disease 2019: retrospective study. BMJ 2020;368:m1091. Doi:10.1136/bmj.m1091
6. Greenhalgh T, Koh GCH, Car J. Covid-19: a remote assessment in primary care. BMJ 2020;368:m1182. Doi:https://doi.org/10.1136/bmj.m1182
7. Chorin E, Padegimas A, Havakuk O, et al. Assessment of respiratory distress by the Roth Score. Clin Cardiol 2019;39(11):636-639.
8. Fletcher CM. Standardised questionnaire on respiratory symptoms: a statement prepared and approved by the MRC Committee on the Aetiology of Chronic Bronchitis (MRC breathlessness score). BMJ 1960;2:1665.

9. Friedman J, Calderón-Villarreal A, Bojorquez I, et al. Excess out-of-hospital mortality and declining oxygen saturation documented by EMS during the COVID-19 crisis in Tijuana, Mexico. medRxiv 2020;e-pub ahead of print. Doi: https://doi.org/10.1101/2020.05.13.20098186.t
10. Ottestad W, Seim M, Maehlen JO. Covid-19 with silent hypoxemia. Tidsskr Nor Legeforen 2020;e-pub ahead of print. Doi:10.4045/tidsskr.20.0299
11. Wilkerson RG, Alder JD, Shah NG, Brown R. Silent hypoxia: A harbinger of clinical deterioration in patients with COVID-19. Am J Emerg Med 2020;e-pub ahead of print. Doi:https://doi.org/10.1016/j.ajem.2020.05.044
12. Gattinoni L, Chiumello D, Caironi P, et al. COVID-19 pneumonia: different respiratory treatments for different phenotypes? Intensive Care Med 2020; e-pub ahead of print. Doi: https://doi.org/10.1007/s00134-020-06033-2
13. Gondos T, Szabo V, Sarkany A, Sarkany A, Halasz G. Estimation of the severity of breathlessness in the emergency department: a dyspnea score. BMC Emergency Medicine 2016;17:13. Doi: 10.1186/s12873-017-0125-6
14. Guryay MS, Ceylan E, Gunay T, et al. Can Spirometry, pulse oximetry and dyspnea scoring reflect respiratory failure in patients with chronic obstructive pulmonary disease exacerbation? Medical Principles and Practice 2007;16:378-383. Doi:10.1159/000104812
15. Higashimoto Y, Honda N, Yamagata T. Exertional dyspnoea and cortical oxygenation in patients with COPD. European Respiratory Journal 2015;46:1615-1624. Doi:10.1183/13993003.00541-2015

16. Tanaka K, Akechi T, Okuyama T. Factors Correlated with Dyspnea in Advanced Lung Cancer Patients: Organic Causes and What Else? Journal of Pain and Symptom Management 2002;23(6):490-500. Doi:10.1016/s0885-3924(02)00400-1

17. Lam PW, Sehgal P, Andany N, et al. A virtual care program for outpatients diagnosed with COVID-19: A feasibility study. CMAJ Open 2020;8(2):E407-413.

18. Manali ED, Lyberopoulos P, Triantafillidou C, et al. MRC chronic dyspnea scale: relationships with cardiopulmonary exercise testing and 6-minute walk test in idiopathic pulmonary fibrosis patients: a prospective study. BMC Pulmonary Medicine 2010;10:32.

19. Clemens KE, Quednau I, Klaschik E. Use of oxygen and opioids in the palliation of dyspnoea in hypoxic and non-hypoxic palliative care patients: a prospective study. Support Care Cancer 2009;17:367-377.

20. Jouffroy R, Jost D, Prunet B. Prehospital pulse oximetry: a red flag for early detection of silent hypoxemia in COVID-19 patients. Critical Care 2020;24:313.

21. Tobin MJ, Laghi F, Jubran A. "Why COVID-19 silent hypoxemia is baffling to physicians. Annals of the American Thoracic Society 2020; e-pub ahead of print. doi:10.1164/rccm.202006-2157CP

22. Gattinoni L, Coppola S, Cressoni M, et al. COVID-19 does not lead to a "typical" acute respiratory distress syndrome. Am J Respir Crit Care Med 2020;201(10):1299-1300.

23. Gattinoni L, Chiumello D, Rossi S. COVID-19 pneumonia: ARDS or not? Critical Care 2020;24:154

24. Low T, Cherian R, Lim SL, et al. Rethinking COVID-19 'pneumonia'-is this primarily a vaso-occlusive disease, and can early anticoagulation save the ventilator famine. Pulmonary Circulation 2020;10(3):1-3.

25. Couzel-Frankel J. The mystery of the pandemic's 'happy hypoxia'. Science 2020;368(6490);455-456.

26. Luks AM, Swenson ER. Pulse oximetry for monitoring patients with COVID-19 at home: potential pitfalls and practical guidance. Annals of the American Thoracic Society 2020; e-pub ahead of print. doi: https://doi.org/10.1513/AnnalsATS.202005-418FR

Figure Legends:

Figure 1. Comparison of SpO2 and measures of subjective dyspnea. A) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients who reported dyspnea and those who did not. B) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients with various mMRC Dyspnea Scale scores. The width of each plot is proportional to the number of patients with the respective SpO2 (represented by black dots). The median SpO2 is indicated by the central horizontal black line and the dotted lines correspond to the interquartile range.

Table 1: Demographics and Clinical Characteristics Among Outpatients with COVID-19 Monitored with Home Oxygen Saturation Devices	
Demographic information	**No. (%)**
Total No.	89
Age, median (IQR), y	52 (38-64)
Sex	
Female	57(64)
Male	32 (36)
Pregnant	6 (7)
Days from symptom onset to clinical assessment, median (IQR)	6 (3-8)
Comorbidities	
Cardiac disease	7 (8)
Chronic lung disease	3 (3)
Asthma	14 (16)
Chronic kidney disease	7 (8)
Moderate/severe liver disease	2 (2)
Chronic neurological issues	5 (6)
Malignancy	14 (16)
Chronic hematological disease	7 (8)
Diabetes	15 (17)
Hypertension	32 (36)
Rheumatic disorder	3 (3)
Malnutrition	1 (1)
Obesity	18 (20)
None	29 (33)
Signs and symptoms on intake assessment	
Fever	35 (39)
Sore Throat	28 (31)
Runny Nose	29 (33)
Cough	56 (63)
Shortness of Breath	23 (26)
Chills/Rigors	30 (34)
Conjunctivitis	10 (11)
Ear Pain	7 (8)
Anosmia	21 (24)
Dysgeusia	25 (28)
Sputum	10 (11)
Hemoptysis	0
Wheezing	7 (8)
Chest Pain	19 (21)
Myalgia	40 (45)
Arthralgia	16 (18)
Abdominal Pain	14 (16)

Nausea/Vomiting	22 (25)
Diarrhea	25 (28)
Adenopathy	0
Rash	1 (1)
Fatigue/Malaise	59 (66)
Headache	37 (42)
Confusion	5 (6)
Depression/Anxiety	15 (17)
Insomnia	19 (21)
Anorexia	33 (37)
Laboratory findings at admission, median (IQR)	
Leukocytes, $x10^9$/L (n=21)	5.9 (4.2-7.5)
Lymphocytes, $x10^9$/L (n=21)	1.1 (0.6-1.3)
Lactate Dehydrogenase, IU/L (n=4)	216.0 (63.4-277.0)
D-dimer, mcg/L (n=5)	906.0 (555.0-1082.5)
High-sensitivity Troponin T ng/L (n=11)	9.7 (6.0-10.0)
Ferritin, mcg/L (n=3)	1644.0 (153.5-2082.5)
Chest radiography done	29 (33)
Abnormal	23 (26)
Bilateral infiltrates	18 (20)
Outcome	
ICU admission	5 (6)
Length of ICU stay, median (IQR), days	6 (2-11)
Intubation	1 (1)
Duration of intubation, days	15 (17)
Hospitalization	11 (12)
Duration of hospitalization, median (IQR), days	3 (2.5-7)
Multiple hospitalizations	2 (2)
Death (within 30 days of diagnosis)	0

Table 2: **Association of Dyspnea Measurements with Detection of Hypoxemia**			
Dyspnea measurement	**Non-hypoxic patients (O2 sat ≥95%)**	**Hypoxic Patients (O2 sat <95%)**	**p-value**
Shortness of breath	14	8	0.046*
Breathing faster at rest	3	0	1.00
Breathing harder than normal	2	0	1.00
More breathless today than yesterday	3	1	0.57
mMRC Dyspnea scale:			
>0	9	7	0.014*
>1	1	5	0.001*
>2	1	5	0.001*
>3	1	1	0.37
Roth Score: Maximum Count			
<12	0	1	0.21
<15	3	2	0.27
<20	5	2	0.60
<28	6	2	1.00
Roth Score: Count time			
<8 sec	1	1	0.39
<15 sec	10	2	0.60
<20 sec	12	2	0.27
<25 sec	13	3	0.53

Table 3: Diagnostic Accuracy of Dyspnea Measurements in the Detection of Hypoxemia: Estimation of test characteristics including sensitivity (SN), specificity, negative (NPV) and positive (PPV) predictive value and negative (-LR) and positive (+LR) likelihood ratio

Dyspnea measurement	SpO2 Cutoff	SN % (95%CI)	SP % (95%CI)	NPV % (95%CI)	PPV % (95%CI)	-LR (95%CI)	+LR (95%CI)
Shortness of breath	≤97%	25 (16-37)	75 (47-92)	19 (10-30)	82 (59-94)	1.0 (0.7-1.4)	1.0 (0.4-2.6)
	≤96%	27 (17-41)	78 (60-90)	39 (27-51)	68 (45-85)	0.9 (0.7-1.2)	1.3 (0.6-2.7)
	≤95%	40 (23-59)	83 (70-91)	72 (60-82)	55 (33-75)	0.7 (0.5-1.0)	2.3 (1.1-4.7)
	≤94%	47 (24-72)	80 (68-88)	86 (75-93)	36 (18-59)	0.7 (0.4-1.1)	2.4 (1.2-4.7)
	≤93%	46 (18-75)	78 (66-86)	91 (80-96)	23 (9-46)	0.7 (0.4-1.2)	2.0 (0.9-4.4)
	≤92%	50 (19-81)	77 (66-85)	95 (86-99)	14 (4-36)	0.6 (0.3-1.5)	2.1 (0.9-5.2)
Breathing faster at rest	≤94%	0 (0-24)	96 (87-99)	81 (70-88)	0 (0-69)	1.0 (1.0-1.1)	0
Breathing harder than normal	≤94%	0 (0-24)	97 (89-100)	81 (71-88)	0 (0-80)	1.0 (1.0-1.1)	0
More breathless today than yesterday	≤94%	6.2 (0.3-32)	96 (87-99)	82 (71-89)	25 (1-78)	1.0 (0.9-1.1)	1.5 (0.2-13.1)
mMRC Dyspnea scale:							
>0	≤94%	54 (26-80)	82 (68-91)	87 (74-95)	44 (21-69)	0.6 (0.3-1.0)	3.0 (1.4-6.5)
>1	≤94%	39 (15-68)	98 (88-100)	86 (74-93)	83 (37-99)	0.6 (0.4-1.0)	19.2 (2.4-151)
>2	≤94%	39 (15-68)	98 (88-100)	86 (74-93)	83 (37-99)	0.6 (0.4-1.0)	19.2 (2.4-151)
>3	≤94%	8 (0.4-38)	98 (88-100)	80 (68-89)	50 (9.5-91)	0.9 (0.8-1.1)	3.9 (0.3-57.4)
Roth Score: Maximum Count							
<12	≤94%	25 (1.3-78)	100 (75-100)	83 (58-96)	100 (6-100)	0.8 (0.4-1.3)	N/A
<15	≤94%	50 (15-85)	80 (51-95)	86 (56-98)	40 (7-83)	0.6 (0.2-1.7)	2.5 (0.6-10.2)
<20	≤94%	50 (15-85)	67 (39-87)	83 (51-97)	29 (5.1-70)	0.8 (0.3-2.1)	1.5 (0.5-5.1)
<28	≤94%	7.7 (0.4-38)	94 (82-98)	79 (66-88)	25 (1.3-78)	1.0 (0.3-2.4)	1.3 (0.4-4.0)
Roth Score: Count time							
<8 sec	≤94%	25 (1.3-78)	93 (66-100)	82 (56-95)	50 (10-91)	0.8 (0.5-1.4)	3.8 (0.3-47.7)
<15 sec	≤94%	50 (15-85)	33 (13-61)	71 (30-95)	17 (2.9-49)	1.5 (0.5-5.1)	0.8 (0.3-2.1)
<20 sec	≤94%	50 (15-85)	20 (5.3-49)	60 (17-93)	14 (2.5-44)	2.5 (0.6-10.2)	0.6 (0.2-1.7)
<25 sec	≤94%	75 (22-99)	13 (2.3-42)	67 (13-98)	19 (5.0-46)	1.9 (0.2-15.8)	0.9 (0.5-1.6)
All Predictors Combined*	≤94%	59 (34-81)	67 (55-78)	87 (75-94)	30 (16-49)	0.6 (0.3-1.1)	1.8 (1.0-3.0)

*This represents a single variable in which all of the measurements were combined into a single predictor: dyspnea OR breathing faster OR breathing harder OR more breathless OR mMRC >0 OR Roth Maximum Count <20 OR Roth Count Time <20

Table 4: Diagnostic accuracy of the presence of dyspnea in the detection of hypoxemia stratified based on patient characteristics, including age, presence vs absence of underlying lung disease, and date from symptom onset

	SN % (95%CI)	SP % (95%CI)	NPV % (95%CI)	PPV % (95%CI)	-LR (95%CI)	+LR (95%CI)
Age						
<60 y	75 (36-96)	70 (54-83)	94 (78-99)	32 (14-57)	0.4 (0.1-1.2)	2.5 (1.3-4.5)
≥60 y	22 (3.9-60)	96 (79-100)	79 (61-90)	67 (13-98)	0.8 (0.6-1.2)	6.0 (0.6-58.6)
Underlying lung disease						
Yes	100 (20-100)	80 (51-95)	100 (70-100)	40 (7.3-83)	0	5.0 (1.8-13.8)
No	40 (18-67)	80 (67-89)	83 (70-92)	35 (15-61)	0.8 (0.5-1.2)	2.0 (0.9-4.5)
Days from symptom onset						
<7 days	50 (19-81)	83 (68-93)	92 (78-98)	30 (8.1-65)	0.6 (0.3-1.4)	3.0 (1.1-8.6)
≥7 days	50 (24-76)	71 (48-88)	75 (51-90)	46 (18-75)	0.7 (0.4-1.4)	1.8 (0.7-4.4)

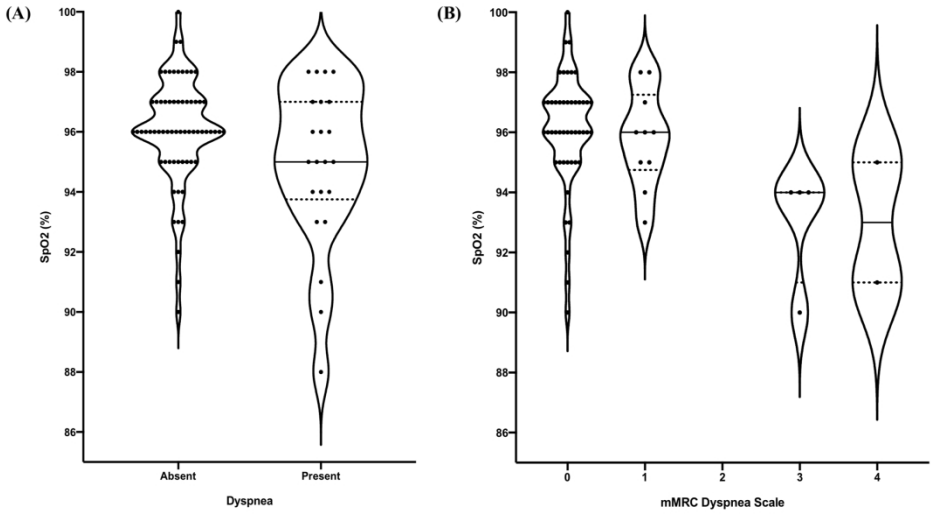

Figure 1. Comparison of SpO2 and measures of subjective dyspnea. A) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients who reported dyspnea and those who did not. B) Violin plots showing the distribution of SpO2 (%) values in COVID-19 outpatients with various mMRC Dyspnea Scale scores. The width of each plot is proportional to the number of patients with the respective SpO2 (represented by black dots). The median SpO2 is indicated by the central horizontal black line and the dotted lines correspond to the interquartile range.

Supplementary Material

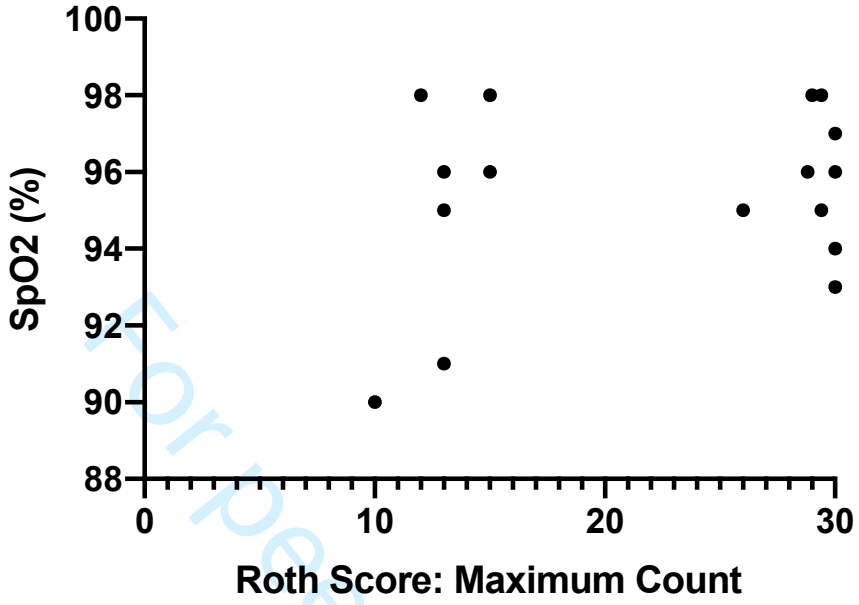

Figure S1. Scatter Plot of SpO2 values across patients' Roth Scores (Maximum Count)

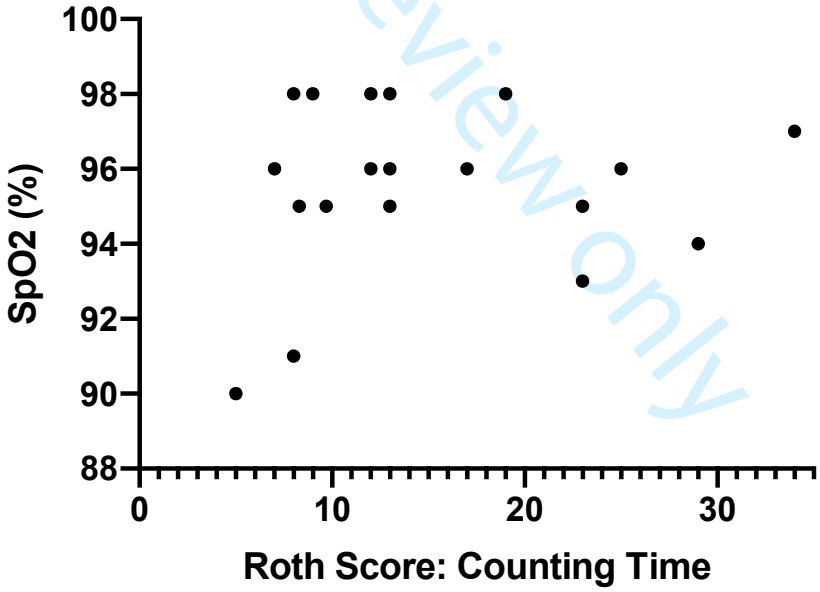

Figure S2. Scatter Plot of SpO2 values across patients' Roth Scores (Counting Time)

Section & Topic	No	Item	Reported on page #
TITLE OR ABSTRACT			
	1	Identification as a study of diagnostic accuracy using at least one measure of accuracy (such as sensitivity, specificity, predictive values, or AUC)	3
ABSTRACT			
	2	Structured summary of study design, methods, results, and conclusions (for specific guidance, see STARD for Abstracts)	3
INTRODUCTION			
	3	Scientific and clinical background, including the intended use and clinical role of the index test	5-6
	4	Study objectives and hypotheses	6
METHODS			
Study design	**5**	Whether data collection was planned before the index test and reference standard were performed (prospective study) or after (retrospective study)	6
Participants	**6**	Eligibility criteria	6
	7	On what basis potentially eligible participants were identified (such as symptoms, results from previous tests, inclusion in registry)	6
	8	Where and when potentially eligible participants were identified (setting, location and dates)	6
	9	Whether participants formed a consecutive, random or convenience series	6
Test methods	**10a**	Index test, in sufficient detail to allow replication	8
	10b	Reference standard, in sufficient detail to allow replication	8
	11	Rationale for choosing the reference standard (if alternatives exist)	8
	12a	Definition of and rationale for test positivity cut-offs or result categories of the index test, distinguishing pre-specified from exploratory	8
	12b	Definition of and rationale for test positivity cut-offs or result categories of the reference standard, distinguishing pre-specified from exploratory	7-8
	13a	Whether clinical information and reference standard results were available to the performers/readers of the index test	7-8
	13b	Whether clinical information and index test results were available to the assessors of the reference standard	7-8
Analysis	**14**	Methods for estimating or comparing measures of diagnostic accuracy	8-9
	15	How indeterminate index test or reference standard results were handled	8-9
	16	How missing data on the index test and reference standard were handled	8-9
	17	Any analyses of variability in diagnostic accuracy, distinguishing pre-specified from exploratory	8-9
	18	Intended sample size and how it was determined	9
RESULTS			
Participants	**19**	Flow of participants, using a diagram	N/A
	20	Baseline demographic and clinical characteristics of participants	9, Table 1
	21a	Distribution of severity of disease in those with the target condition	10
	21b	Distribution of alternative diagnoses in those without the target condition	N/A
	22	Time interval and any clinical interventions between index test and reference standard	N/A
Test results	**23**	Cross tabulation of the index test results (or their distribution) by the results of the reference standard	11-12, Table 3
	24	Estimates of diagnostic accuracy and their precision (such as 95% confidence intervals)	11-12, Table 3
	25	Any adverse events from performing the index test or the reference standard	N/A
DISCUSSION			
	26	Study limitations, including sources of potential bias, statistical uncertainty, and generalisability	14
	27	Implications for practice, including the intended use and clinical role of the index test	15
OTHER INFORMATION			
	28	Registration number and name of registry	N/A
	29	Where the full study protocol can be accessed	6
	30	Sources of funding and other support; role of funders	N/A

STARD 2015

AIM

STARD stands for "Standards for Reporting Diagnostic accuracy studies". This list of items was developed to contribute to the completeness and transparency of reporting of diagnostic accuracy studies. Authors can use the list to write informative study reports. Editors and peer-reviewers can use it to evaluate whether the information has been included in manuscripts submitted for publication.

EXPLANATION

A **diagnostic accuracy study** evaluates the ability of one or more medical tests to correctly classify study participants as having a **target condition.** This can be a disease, a disease stage, response or benefit from therapy, or an event or condition in the future. A medical test can be an imaging procedure, a laboratory test, elements from history and physical examination, a combination of these, or any other method for collecting information about the current health status of a patient.

The test whose accuracy is evaluated is called **index test.** A study can evaluate the accuracy of one or more index tests. Evaluating the ability of a medical test to correctly classify patients is typically done by comparing the distribution of the index test results with those of the **reference standard**. The reference standard is the best available method for establishing the presence or absence of the target condition. An accuracy study can rely on one or more reference standards.

If test results are categorized as either positive or negative, the cross tabulation of the index test results against those of the reference standard can be used to estimate the **sensitivity** of the index test (the proportion of participants *with* the target condition who have a positive index test), and its **specificity** (the proportion *without* the target condition who have a negative index test). From this cross tabulation (sometimes referred to as the contingency or "2x2" table), several other accuracy statistics can be estimated, such as the positive and negative **predictive values** of the test. Confidence intervals around estimates of accuracy can then be calculated to quantify the statistical **precision** of the measurements.

If the index test results can take more than two values, categorization of test results as positive or negative requires a **test positivity cut-off**. When multiple such cut-offs can be defined, authors can report a receiver operating characteristic (ROC) curve which graphically represents the combination of sensitivity and specificity for each possible test positivity cut-off. The **area under the ROC curve** informs in a single numerical value about the overall diagnostic accuracy of the index test.

The **intended use** of a medical test can be diagnosis, screening, staging, monitoring, surveillance, prediction or prognosis. The **clinical role** of a test explains its position relative to existing tests in the clinical pathway. A replacement test, for example, replaces an existing test. A triage test is used before an existing test; an add-on test is used after an existing test.

Besides diagnostic accuracy, several other outcomes and statistics may be relevant in the evaluation of medical tests. Medical tests can also be used to classify patients for purposes other than diagnosis, such as staging or prognosis. The STARD list was not explicitly developed for these other outcomes, statistics, and study types, although most STARD items would still apply.

DEVELOPMENT

This STARD list was released in 2015. The 30 items were identified by an international expert group of methodologists, researchers, and editors. The guiding principle in the development of STARD was to select items that, when reported, would help readers to judge the potential for bias in the study, to appraise the applicability of the study findings and the validity of conclusions and recommendations. The list represents an update of the first version, which was published in 2003.

More information can be found on http://www.equator-network.org/reporting-guidelines/stard.