UC San Diego

Title

Breath-holding physiology, radiological severity and adverse outcomes in COVID-19 patients: A prospective validation study

Permalink

https://escholarship.org/uc/item/9gh8z0mz

Journal

Respirology, 27(12)

ISSN

1323-7799

Authors

Messineo, Ludovico
Fanfulla, Francesco
Pedroni, Leonardo
et al.

Publication Date

2022-12-01

DOI

10.1111/resp.14336

Peer reviewed

ORIGINAL ARTICLE

Breath-holding physiology, radiological severity and adverse outcomes in COVID-19 patients: A prospective validation study

Ludovico Messineo[1] | Francesco Fanfulla[2] | Leonardo Pedroni[3] | Floriana Pini[4] |
Andrea Borghesi[5] | Salvatore Golemi[5] | Guido Vailati[3] | Kayla Kerlin[1] |
Atul Malhotra[6] | Luciano Corda[3,7] | Scott Sands[1,8]

[1]Division of Sleep and Circadian Disorders, Departments of Medicine and Neurology, Brigham & Women's Hospital & Harvard Medical School, Boston, Massachusetts, USA

[2]Sleep Medicine and Respiratory Function Unit, Maugeri Clinical and Scientific Institutes IRCCS, Pavia, Italy

[3]Respiratory Medicine and Sleep Laboratory, Department of Experimental and Clinical Sciences, University of Brescia and Spedali Civili, Brescia, Italy

[4]Department of General Medicine, Mellini Hospital, Italy

[5]Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, University of Brescia, Brescia, Italy

[6]Division of Pulmonary and Critical Care Medicine, University of California San Diego, La Jolla, California, USA

[7]Department of Internal Medicine, Spedali Civili, Brescia, Italy

[8]Department of Allergy Immunology and Respiratory Medicine and Central Clinical School, The Alfred and Monash University, Melbourne, Victoria, Australia

Correspondence
Ludovico Messineo
Email: lmessineo@bwh.harvard.edu

Associate Editor: Alexander N Larcombe; **Senior Editor:** Phan Nguyen

Abstract

Background and objective: COVID-19 remains a major cause of respiratory failure, and means to identify future deterioration is needed. We recently developed a prediction score based on breath-holding manoeuvres (desaturation and maximal duration) to predict incident adverse COVID-19 outcomes. Here we prospectively validated our breath-holding prediction score in COVID-19 patients, and assessed associations with radiological scores of pulmonary involvement.

Methods: Hospitalized COVID-19 patients ($N = 110$, three recruitment centres) performed breath-holds at admission to provide a prediction score (Messineo et al.) based on *mean desaturation* (20-s breath-holds) and *maximal breath-hold duration*, plus baseline saturation, body mass index and cardiovascular disease. Odds ratios for incident adverse outcomes (composite of bi-level ventilatory support, ICU admission and death) were described for patients with versus without elevated scores (>0). Regression examined associations with chest x-ray (*Brixia* score) and computed tomography (CT; 3D-software quantification). Additional comparisons were made with the previously-validated '4C-score'.

Results: Elevated prediction score was associated with adverse COVID-19 outcomes ($N = 12/110$), OR[95%CI] = 4.54[1.17–17.83], $p = 0.030$ (positive predictive value = 9/48, negative predictive value = 59/62). Results were diminished with removal of *mean desaturation* from the prediction score (OR = 3.30[0.93–11.72]). The prediction score rose linearly with *Brixia* score (β[95%CI] = 0.13[0.02–0.23], $p = 0.026$, $N = 103$) and CT-based quantification ($\beta = 1.02[0.39–1.65]$, $p = 0.002$, $N = 45$). *Mean desaturation* was also associated with both radiological assessment. Elevated 4C-scores ($\geq$high-risk category) had a weaker association with adverse outcomes (OR = 2.44[0.62–9.56]).

Conclusion: An elevated breath-holding prediction score is associated with almost five-fold increased adverse COVID-19 outcome risk, and with pulmonary deficits observed in chest imaging. Breath-holding may identify COVID-19 patients at risk of future respiratory failure.

KEYWORDS
breath-holding procedure, coronavirus disease, COVID-19, incident adverse outcome, mean desaturation, physiology, radiological severity

INTRODUCTION

Since the onset of the COVID-19 pandemic, a major challenge for clinicians has been to interpret which patients who present with COVID-19 are most likely to exhibit future respiratory failure and which patients are likely to recover with minimal intervention. Two years into the pandemic, nearly 10% of unvaccinated individuals with COVID-19 are hospitalized,[1,2] approximately half of whom later develop respiratory failure and require intensive care.[3] These data are of particular concern for regions with reduced vaccination access or uptake.[4–12] During each of the

surges in COVID case rates, hospital resources have been under immense pressure.[13] Physicians are asked to decide which patients are at greatest risk for respiratory failure for early administration of finite supplies of novel therapeutics or prioritization of limited intensive care resources.[14,15] Thus, there remains a need for practicable triage strategies to facilitate recognition of cases at highest risk of deterioration.

Recently, utilizing a previously-validated technique,[16] we demonstrated that simple physiological measures of gas exchange and neurophysiological (i.e., chemosensitivity) deficits obtained using breath-holding manoeuvres were associated with adverse outcomes of COVID-19.[17] Additionally, we proposed that breath-holding-related desaturation and maximal breath-duration, in addition to baseline saturation, body-mass index and cardiovascular disease status, could serve as a potential means to predict susceptibility to COVID-19 incident adverse outcome (i.e., respiratory failure and/or death).[17] In principle, a greater ease of desaturation with breath-holding was considered to unmask V–Q heterogeneity and reduced lung volume that may not be well-captured by baseline oxygen saturation levels. In addition, a greater tolerance for breath holding (increased maximal duration adjusted for hypoxaemia) was considered to reflect 'silent hypoxaemia' (disproportionate tolerance of hypoxaemia), a likely risk factor for further progression of infection.[18] However, as highlighted by an ongoing review of COVID-19 prediction models,[19] our tool is at high risk of bias due to lack of external validation.

Accordingly, in a pragmatic, multi-centre, observational study, we sought to validate prospectively whether our prediction model that incorporates markers of gas exchange and chemosensitivity deficits predicts increased odds of adverse COVID-19 outcomes (composite of bi-level ventilatory support, intensive care unit [ICU] admission and death). Additionally, we aimed to seek associations between preferential decrements in our model's prediction score and increased severity of COVID-19 pneumonia seen on chest x-ray and/or computed tomography (CT). Comparisons were also made against a previously-validated, biomarker-based COVID-19 prediction score ('4C-score').[20]

METHODS

Participant recruitment

110 hospitalized patients aged 18–90 years, positively swabbed for COVID-19, were enrolled in three different centres in northern Italy (Brescia, Chiari and Pavia). Exclusion criteria were: more-than-moderate dyspnoea (Borg ≥ 6), hemodynamic instability, Brescia-COVID respiratory severity scale >1,[21] diurnal home treatment with supplemental oxygen or ventilatory support, use of sedatives, opioids, anti-emetics or other drugs known to impact chemosensitivity, heart failure, chronic obstructive pulmonary disease, pregnancy and inability to understand the informed consent. Patients were recruited from December 2020 until July 2021.[22]

SUMMARY AT A GLANCE

An elevated breath-holding-based prediction score was associated with increased COVID-19 incident adverse outcome risk in a validation cohort of 110 hospitalized COVID-19 patients. The prediction score was also positively associated with increasing radiological severity, per chest x-ray and computed tomography (CT) assessment. Our prediction score performed better than the previously-validated, biomarker-based 4C-score.

Study protocol

For each Covid-19 patient, the following data were collected on admission to the Covid-19 dedicated unit: medical history, medications, baseline demographics (i.e., age, body mass index [BMI] and ethnicity), baseline peripheral oxy-haemoglobin saturation (SpO_2), heart rate and blood pressure, Glasgow Coma Scale, presence of dyspnoea (continuous and categorical: Borg = 1–5 vs. Borg = 0) or other current symptoms of Covid-19 (anosmia, ageusia, gastro-intestinal symptoms, yes/no, and history of smoking [current, prior and never]). Blood tests included measurement of C-reactive protein (CRP) levels, white blood cell count and formula, haemoglobin and urea. Subsequently, breath-holds were performed (details below). Per clinical indication, a chest x-ray and/or a chest CT scan was performed. The 4C-score was also quantified for each patient based on eight recorded parameters[20]: age, sex at birth, number of comorbidities, respiratory rate, oxygen saturation, Glasgow Coma Scale and urea and CRP levels; a score of 9 or more was considered high risk.

Patients were then treated according to standard medical care, regulated by national protocols of the Italian Health Care System. The following outcomes, if present, were recorded: first administration of oxygen therapy, first administration of bi-level PAP, admission to ICU, death or discharge.

Breath-holding procedure

Patients performed breath-holds shortly after admission and while breathing room air. Supplemental oxygen, when needed ($N = 77$), was temporarily disconnected for the duration of the procedure (~15 min).

Breath-holding manoeuvres were performed as previously described.[16] Briefly, while supine, participants were instructed to hold their breath starting from functional residual capacity (FRC) and avoid deep inspiration prior to breath-holds. Investigators requested at least four 20-s fixed-time breath-holds and at least 1 maximal breath-hold. An investigator sat at bedside and described when to start each

breath-hold while placing a hand on the patients' lower chest. Verbal encouragement was also provided during maximal breath-holds.

Unlike prior detailed physiology studies,[16,17] airflow and oxygen saturation signals were not recorded. For clinical applicability, visual observation of oximetre values (through probes provided by the hospital, with an averaging time of 12-s or faster, for example, *8000S*, Nonin, Plymouth, MN; *Tuffsat*, GE Datex, Chicago, IL; *iPM8*, Mindray, Shenzen, CN) was used to record *baseline saturation* prior to breath-holds and the greatest single nadir desaturation following the 20-s breath-holds (a simplified, pragmatic estimate of the *mean desaturation* variable that we described previously[17]). *Maximal breath-hold duration* (largest value observed) was recorded via a timing device.

Chest imaging

Available chest x-ray and/or CT data were evaluated in consensus by two thoracic radiologists (Andrea Borghesi, Salvatore Golemi). The radiological assessment of the severity of COVID-19 pneumonia was ranked via: (1) *Brixia* score,[23] a dedicated chest x-ray scoring system where each pulmonary zone (upper, middle and lower in both lungs and frontal chest projection) was scored (0–3 points) based on the severity of lung abnormalities (score range: 0–18); (2) Computer-aided CT-based quantification of percentage of pulmonary involvement (quantification range: 0%–100%) was performed by applying a dedicated three-dimensional software (Syngo.via, Siemens Healthcare).[24] In addition, chest CT examinations were also scored by applying the *Brixia* score (score range: 0–18) on a coronal multiplanar reformation image (c-MPR) reconstructed to resemble a frontal chest x-ray.[25]

Statistical analysis

The adverse composite outcome was reached with any occurrence of non-invasive bi-level pressure support, intensive care admission or death. Herein, we refer to patients that met the adverse outcome as VS+ ('ventilatory support'), while patients discharged without experiencing the adverse outcome were labelled VS−.

Our a priori hypotheses and study outcomes were pre-specified in a statistical analysis plan. The outcome database was locked prior to data analysis.

The primary study analysis involved comparing the actual occurrence of the adverse outcome versus the predicted adverse outcome provided by our prediction score. Briefly, *mean desaturation* and *maximal breath hold duration* were combined with covariates (baseline saturation, BMI and cardiovascular disease) into a model that indicates elevated risk of adverse outcomes with scores above 0 (i.e., expected >50% likelihood of adverse outcome). Odds ratios (Fisher exact-test) for incident adverse COVID-19

outcomes were described for patients with versus without elevated prediction scores. In exploratory analysis, we evaluated the separate role of each of the breath-holding measures in the primary outcome analysis by removing the data for the particular measure (mean value replacement) and repeating the odds ratio assessment. Positive and negative predictive values were also described.

Sensitivity analysis examined the predictive value of the continuous breath-holding prediction score using logistic regression; receiver operator characteristic area under the curve (AUC) was described.

Linear regression was used to determine whether the prediction score was associated with radiological severity assessed via chest x-ray (per *Brixia* score, continuous) and chest CT (per computer-aided quantification, continuous). Further analysis tested potential association between the prediction score and the *Brixia* score applied to c-MPR CT images. Mechanistic analysis assessed whether *mean desaturation* specifically (unadjusted or adjusted for baseline saturation, BMI, cardiovascular disease, sex and age) was associated with these radiological severity scores, as expected if breath-holding desaturation is indeed a biomarker of pulmonary deficits (low lung gas volume, ventilation-perfusion heterogeneity); the potential confounding effect of pre-admission therapy with antibiotics, steroids and vaccination status on the relationship between *mean desaturation* and the radiological severity scores was assessed with separate multiple linear regressions.

We also assessed whether the breath-holding prediction score was associated with the currently-available 4C-score, using dichotomous analysis (breath-holding prediction score > 0 vs. 4C-score ≥ 9; odds ratio per Fisher exact test) or simple linear regression on the continuous prediction scores. Logistic regression also assessed whether mean desaturation and/or maximal breath-hold duration were associated with the 4C-score (adjusted for covariates above, or separately for pre-admission treatment with antibiotics and/or steroids or vaccination status). Finally, we examined the predictive value of the 4C-score for adverse outcomes of COVID-19 in our population (odds ratio per Fisher exact test).

A sample size of 115 (105 complete plus 10 dropouts) was estimated a priori based to provide a 90% power to detect an odds ratio of ~6 and 80% power to detect an odds ratio of ~5 (alpha = 0.05, ratio of favourable to unfavourable test prognosis 4:1).[17] A *p* value of <0.05 was considered statistically significant. Statistical analysis was performed with the MATLAB (Mathworks, Natick, MA).

RESULTS

One hundred and ten patients were enrolled and all individuals provided data for analysis. Patient baseline characteristics are presented in Table 1. N = 12/110 patients with COVID-19 met the primary composite outcome criteria (ventilatory support, intensive care or death). Chest x-rays were performed in N = 103, and chest CT were performed

T A B L E 1 Baseline characteristics

Characteristic	All ($N = 110$)	VS+ ($N = 12$)	VS− ($N = 98$)
Population factors			
Age, y	63.9 ± 15.4	66.8 ± 11.9	63.6 ± 15.8
Male sex, n (%)	71 (65)	6 (50)	65 (66)
Body mass index, kg/m^2	27.9 ± 5.3	29.2 ± 5.3	27.7 ± 5.3
Caucasian or white race/ethnicity, n (%)	101 (92)	11 (92)	90 (92)
History			
History of hypertension, n (%)	64 (58)	6 (50)	58 (59)
History of cardiovascular disease, n (%)	19 (17)	4 (33)	15 (15)
History of diabetes, n (%)	22 (20)	3 (20)	19 (19)
Current smoking, n (%)	14 (13)	1 (8)	13 (13)
Current medications			
β-blockers, n (%)	32 (29)	5 (42)	27 (28)
ACE-inhibitors, n (%)	23 (21)	0 (0)	23 (23)
ARB, n (%)	18 (16)	2 (17)	16 (16)
COVID-19 therapy before hospital admission			
Antibiotics, n (%)	45 (41)	6 (50)	39 (40)
Steroids, n (%)	31 (28)	6 (50)	25 (26)
Vaccine (one dose), n (%)	7 (6)	1 (8)	6 (6)
Clinical presentation at admission			
Baseline SpO$_2$, %	94.2 ± 3.0	93.3 ± 1.9	94.3 ± 3.1
Baseline heart rate, beats/min	84.5 ± 16.7	85.3 ± 11.8	84.4 ± 17.2
Baseline respiratory rate, breath/min	21.6 ± 6.1	21.2 ± 5.0	21.7 ± 6.2
Baseline systolic blood pressure, mm Hg	132.0 ± 19.6	134.7 ± 19.0	131.6 ± 19.8
Baseline diastolic blood pressure, mm Hg	75.3 ± 13.0	73.3 ± 15.0	75.5 ± 12.8
Glasgow Coma Scale	14.9 ± 0.4	15.0 ± 0	14.9 ± 0.4
Anosmia, n (%)	38 (35)	2 (17)	36 (37)
Ageusia, n (%)	46 (42)	3 (25)	43 (44)
Gastrointestinal symptoms, n (%)	60 (55)	6 (50)	54 (55)
Dyspnoea (Borg)	2 ± 2	2 ± 2	2 ± 2
Laboratory tests			
C-reactive protein, mg/L	55.1 ± 56.8	36.9 ± 29.8	57.2 ± 58.8
White blood count, $n * 10^3$/L	6.1 ± 2.8	4.8 ± 2.4	6.3 ± 2.8
Haemoglobin, g/dl	13.2 ± 2.1	13.1 ± 2.6	13.2 ± 2.0
Urea, mmoL/L	8.0 ± 9.3	9.3 ± 12.8	7.8 ± 8.9
Breath-holding measurements			
Mean desaturation, %	4.8 ± 2.7	6.0 ± 3.4	4.6 ± 2.5
Maximal breath-hold duration, s	27.1 ± 7.4	26.7 ± 6.5	27.2 ± 7.5
Outcomes			
None:oxygen:ventilatory support, n	33:77:12	0:12:12	31:67:0
Duration of hospitalization, d	15.0 ± 17.1	24.3 ± 18.4	14.0 ± 16.7

Note: Data are expressed as mean $\pm$ SD or as N (%). Patients who met the criteria for the adverse primary composite outcome are denoted 'VS+' ($N = 9$ non-invasive bi-level pressure support, $N = 4$ intensive care and $N = 2$ death). In all patients, diagnosis was confirmed with a positive nasal or pharyngeal swab. All patients admitted to ICU were administered mechanical ventilatory support; the patients who died were ventilated with non-invasive bi-level pressure support. 'VS−' indicates patients discharged without meeting adverse composite outcome. Average time to the primary outcome in VS+ patients was 1 ± 2 days. Non-Caucasian/non-White race/ethnicities were black (1 VS+ and 2 VS− COVID-19 patients), Hispanic (1 VS− COVID-19 patients) and Asian (5 VS− COVID-19 patients). 'None' indicates discharge without oxygen or interventions that met criteria for the primary outcome during the hospital stay. On average, 3.5 ± 0.6 20-s breath-holds and 1.9 ± 0.8 maximal breath-holds per individual were performed.

Abbreviations: ACE, angiotensin converting enzyme; ARB, angiotensin II receptor blockers; PaCO$_2$, partial pressure of carbon dioxide; PaO$_2$, partial pressure of oxygen; SpO$_2$, peripheral oxyhaemoglobin saturation.

in $N = 45$. Importantly, there was no adverse events related to breath-holding examination.

Prediction score validation

Elevated prediction score (greater than zero; dichotomous predictor) was associated with adverse COVID-19 outcome, OR[95%CI] = 4.54[1.17–17.83], $p = 0.030$ (positive predictive value = 9/48, negative predictive value = 59/62; Figure 1). In exploratory analysis, this association was not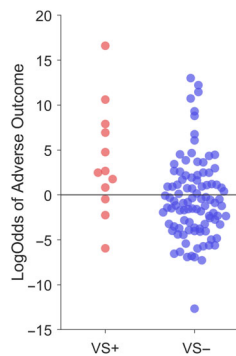 significant anymore with the removal of *mean desaturation* from the prediction score (OR = 3.30[0.93–11.72], $p = 0.067$), suggesting it contributed meaningfully to the association. However, the removal of the *maximal breath-holding duration* from the prediction had no meaningful impact on the association (OR = 4.73[1.21–18.61], $p = 0.028$). Sensitivity analysis also revealed that a higher continuous prediction score was linearly associated with greater risk of adverse outcome (OR = 5.05[1.57–16.19] per 10-point increase in score, $p = 0.006$) and meaningfully discriminated between individuals with versus without adverse outcomes (AUC = 0.73, sensitivity 75% and specificity 60% for cut-off = 0 and $p = 0.009$).

FIGURE 1 Our prediction score (LogOdds of Adverse Outcome) based on simplified breath-holding manoeuvres discriminates patients with elevated risk of COVID-19 adverse outcome with 76% sensitivity and 60% specificity for a threshold ∼0 (i.e., 50% probability of adverse outcome). Findings are similar to our previous internal validation (see Messineo et al.,[17] Figure 2D). VS+ (ventilatory support), patients that met the composite adverse outcome. VS−, patients that did not meet the composite adverse outcome

Prediction score and radiological severity

The prediction score was linearly associated with radiological severity, per increased *Brixia* score, as applied to both chest x-ray (β[95%CI] = 0.13[0.02–0.23] points per score unit, $p = 0.026$; Figure 2A) and c-MPR CT images ($\beta = 0.24$[0.06–0.41] points per score unit, $p = 0.009$; Figure 2B), and per computer-aided CT-based quantification ($\beta = 1.02$[0.39–1.65] %lung involvement per score unit, $p = 0.002$; Figure 2C). *Mean desaturation* specifically was associated with the *Brixia* score applied to c-MPR CT images ($\beta = 0.49$[0.16–0.82] score units per %desaturation, $p = 0.004$), with computer-aided CT-based quantification ($\beta = 2.16$[0.99–3.34] %lung involvement per %desaturation, $p = 0.001$), and borderline associated with the *Brixia* score assigned on chest x-ray images ($\beta = 0.20$[−0.005–0.41] score units per %desaturation, $p = 0.056$) after adjusting for covariates. Separately adjusting the analysis for treatment with antibiotics or steroids before hospital admission, or
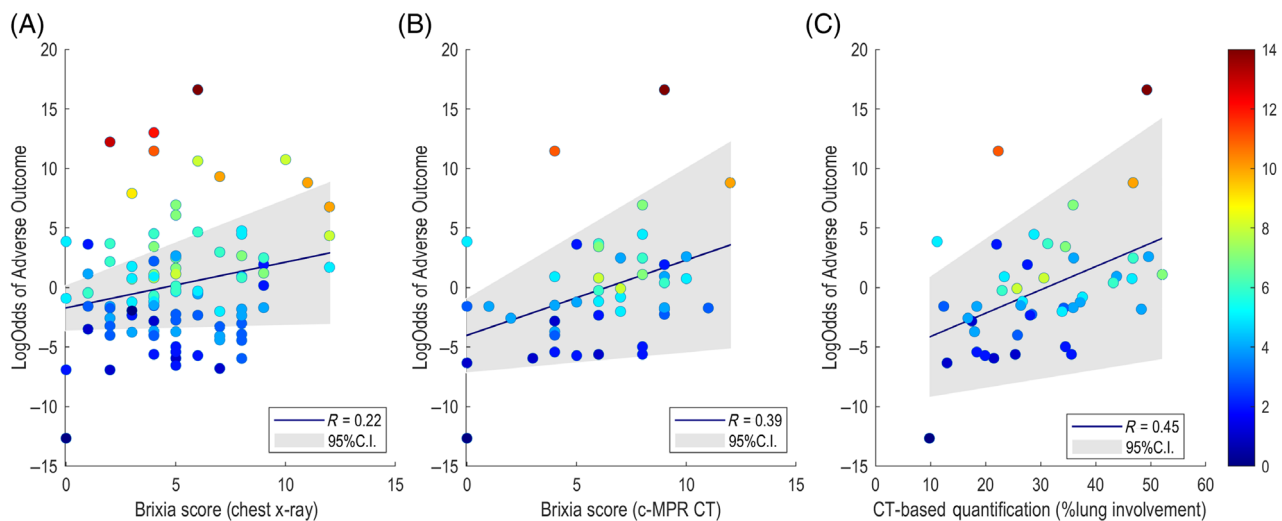

FIGURE 2 The breath-holding based prediction score (LogOdds of Adverse Outcome) was positively associated with the *Brixia* score applied to chest x-ray (panel A; $N = 103$) and applied to c-MPR CT (panel B; $N = 45$) and was associated with computer-aided CT-based quantification of pulmonary involvement (panel C; $N = 45$). Different colours illustrate the magnitude of mean desaturation (i.e., red and blue correspond to large and small values, respectively, as represented by the colour bar). Note that those with greater desaturation, who had higher predicted risk of adverse outcomes (orange-red dots concentrated at the top of the figures, blue at the bottom), tended to exhibit greater radiological severity scores (orange-red dots concentrated to the right of the figures, blue to the left), particularly for CT (panels B,C). c-MPR CT, coronal multiplanar reformation image computed tomography

vaccination rate only slightly changed these findings: *mean desaturation* was significantly associated with *Brixia* score on chest x-ray ($\beta = 0.22[0.02–0.42]$ score units per %desaturation, $p = 0.028$) and computer-aided CT-based, and borderline associated with the *Brixia* score applied to c-MPR CT images ($\beta = 0.34[-0.003–0.679]$ score units per % desaturation, $p = 0.052$). *Maximal breath-hold duration* alone was not associated with any radiological score.

Prediction score and 4C-score

The 4C-score was quantified for all 110 patients. Elevated 4C-score (9 or above; dichotomous predictor) was associated with an elevated breath-holding prediction score ($OR = 4.66[2.01–10.80]$, $p < 0.001$). There was also a linear relationship between the two prediction scores ($\beta = 0.27$ $[0.16–0.38]$ units per score, $p < 0.001$). Again, *mean desaturation*, fully adjusted for covariates ($OR = 1.65[1.01–2.70]$, $p = 0.045$) or for pre-admission therapy and vaccination status ($OR = 1.32[1.13–1.69]$, $p = 0.002$), was associated with an elevated 4C-score. As with the above analysis, the *maximal breath-hold duration* was not associated with an elevated 4C-score. Analysis of the risk associated with an elevated 4C-score yielded a weaker association with the adverse composite outcome in our population ($OR = 2.44$ $[0.62–9.56]$).

DISCUSSION

In this pragmatic observational study, we prospectively-tested the prognostic value of breath-holding prediction score to identify increased risk of adverse outcomes in hospitalized patient with COVID-19. We demonstrated that elevated scores were associated with almost five-fold odds of respiratory failure in our validation cohort of 110 COVID-19 patients, which confirms the cross-validated results of our previous study.[17] The degree of oxygen desaturation that accompanied 20-s-timed breath holds was confirmed as a meaningful contributor to the increased risk. The prediction score was also positively associated with increased chest radiological severity; again, breath-holding desaturation was associated with radiological severity. Finally we also showed that elevated risk per breath-holding prediction score was associated with a previously-validated prediction score (4C-score) based on eight clinical measurements and biomarkers. As opposed to what required to provide the 4C-score, the breath-holding assessment only requires an oximetre and a timing device and is rapidly performed (completion time is ~15 min). If further validated in larger population samples, we envisage our breath-holding prediction score to become a crucial clinical step for rapid triage of COVID-19 patients at high risk of respiratory failure, especially in areas with sustained community transmission and/or low vaccination rate, and where access to other tools for outcome prediction are scarce.

Desaturation

The breath-holding prediction score combined deficits observed in gas exchange (rapid desaturation) and ventilatory control (hyposensitive chemoreflexes per greater breath-holding tolerance) to predict COVID-19 adverse outcomes. Notably, *mean desaturation*, a putative physiology marker of latent V/Q mismatch[26] and reduced lung gas volumes (greater decline in alveolar PO_2 per unit time),[27,28] was a meaningful contributor to the prediction model (odds lowered from five-fold to three-fold without this measurement). Our study further supports the notion that breath-hold induced desaturation can provide meaningful information on gas exchange deficits in COVID-19 beyond baseline saturation alone (which could be insensitive to reduced partial pressure of oxygen [PaO_2] when on the plateau of the SpO_2/PaO_2 curve and thus less prone to capture potential future deterioration). Patients who later deteriorate to respiratory failure therefore have early gas exchange deficits (regions of low V/Q ratio or lower lung gas volume) that are likely a reflection of undetected infection-related deterioration. To our knowledge, an alternative model that includes *dynamic* desaturation for COVID-19 outcome prediction has not been proposed or validated yet. Exertional tests have been widely used for triage early in the pandemic[29,30]; however, their utility for COVID-19 prognostication is still undetermined[31] or questionable.[32] In addition, exertional tests require walking, with increased energy expenditure or cardiac output, and potential contamination of staff/caregivers.

Ventilatory control

Considerable evidence in healthy individuals indicates that a greater *maximal breath-hold duration* reflects a less sensitive ventilatory control system, likely a product of a lower ventilatory chemosensitivity to hypercapnia.[16,18,33] In our prior study of breath-holding physiology in patients with COVID-19, we demonstrated that individuals who did not later develop respiratory failure—as opposed to those who did—had shorter breath-hold durations when adjusting for baseline and breath-hold related hypoxaemia.[17] Thus, the maximal breath-hold duration was included in the *breath-hold* prediction score that was under prospective investigation in the current study. Here, however, we did not find evidence that the prediction score relied on the *maximal breath-hold duration* data. Specifically, removal of this variable did not reduce the odds ratio describing the risk of adverse outcomes (odds ratio remained ~5). Thus, the contribution of *maximal breath-hold duration* to COVID-19 prognostication seems less important here than in the development study. Reasons for this discrepancy are unclear, but might rely on slightly different techniques of breath-holding assessment. In the original study, breath-hold data were recorded with data acquisition software, the test was repeated until obtaining at least two reliable maximal breath-holds

(i.e., precise start at FRC and flat flow throughout the entire manoeuvre) and maximal-breath hold duration was calculated objectively. Therefore, the patients had more trials to understand and adapt to the manoeuvre,[34] and the average breath-hold duration was longer (48.8 ± 16.9 s, compare with Table 1; of note, average desaturation was similar between the studies), suggesting patients may have terminated breath-holds earlier in the current study for reasons unrelated to ventilatory control physiology. In this work, clinical applicability was prioritized to minimize the time and resources required for data collection.

Radiological severity

The findings that our prediction score, and mean desaturation, were positively associated with increased radiological severity provided novel physiological evidence that faster breath-holding desaturation in COVID-19 is indeed a physiological reflection of visually observable pulmonary deficits. Chest x-ray and chest CT are the most commonly used radiological examinations for the management of COVID-19 patients. In addition, both chest x-ray and CT have been shown to predict negative outcomes in COVID-19.[23,24,35] Although chest CT is currently not recommended for initial screening in COVID-19 patients presenting at the emergency department due to high radiation dosage,[36] its role is fundamental in discriminating uncertain cases[37,38] and seeking possible medical complications.[39] We showed that an elevated prediction score highly correlates with CT-based quantification of pulmonary involvement. *Mean desaturation* alone, but not *maximal breath-hold duration*, was also associated with the computer-aided CT-base quantification. Consistent with this finding, ground-glass opacities and consolidated areas of the lungs on CT have been seen prior to respiratory failure[40–42] and likely illustrate compromised regional loss of ventilation and gas volume that we observed in the form of faster desaturation. Our prediction score and *mean desaturation* were also correlated with the *Brixia* score,[23] but *mean desaturation* was only borderline associated with it. However, the *Brixia* score used to assess c-MPR CT images (i.e., images similar to a frontal chest x-ray) was strongly associated with *mean desaturation*. Although less sensitive than chest CT,[43] chest x-ray is typically used for rapid screening in suspected COVID-19 and characteristic findings could also be ground-glass and consolidation areas.[43]

Comparison with the 4C-score

The 4C-score contains parameters reflecting patient demographics, comorbidity, physiology (including baseline $SpO_2 < 92\%$ as a risk factor) and inflammation at hospital admission, and performed well in external validation, unlike similar existing prognostic scores that only performed at best moderately, either because external validation populations were small,[44] and/or diverse in case-mix and severity.[45] The 4C-score was associated with our prediction score, but had weaker association with adverse outcome than our prediction score in our population. Explanations could be that (1) the 4C-score mainly predicts in-hospital mortality risk, and (2) our population sample was not large enough to produce a significant finding. Regardless, our prediction score performed well in our external validation sample and outperformed the 4C-score by magnitude of risk. Further research is warranted to confirm these findings.

Methodological considerations

This study has several limitations. First, to prioritize clinical feasibility, breath-holding manoeuvres were performed slightly differently from the parent study (see Section 2) to obviate the dependence of the measurements on complex data acquisition equipment; however, the fact that the findings were confirmed overall supports the strength of our initial physiological hypothesis and of our prediction model. Second, we did not perform virus sequencing to understand what virus variant the patients had at the time of hospitalization; however, B.1.1.7 (Alpha) was by far the dominant variant in Italy during the study time, with only a peak of B.1.617.2 (Delta) towards the end of the study,[46] while B.1.1.529 (Omicron) and its subvariants were not yet present. It is possible that our results may be less applicable to novel variants (e.g., Omicron is currently considered to affect lung parenchyma less frequently than earlier variants), yet despite a lower hospitalization rate there remains a substantial proportion of individuals requiring advanced care once hospitalized.[47] To the extent that breath-hold induced desaturation is a measure of pulmonary deficits, and pulmonary deficits are responsible for respiratory failure and intensive care requirements, we suspect that associations will continue to be relevant in the future.

Finally, breath-holding required cooperation, so patients who required immediate ventilatory support were not studied, however they were readily triaged and thus beyond any requirement for prediction. Notably, we observe that, compared to the parent study, there was a lower prevalence of patients needing intensive care and a shorter hospitalization stay; this could be the expression of better COVID-19 management and improved therapy since the development study.

In a pragmatic prospective observational study, we validated a prediction model that captures deficits in breath-holding physiology and detects hospitalized patients who are at increased risk of later developing adverse COVID-19 outcomes. An elevated prediction score was associated with almost a five-fold increased odds of respiratory failure, and was associated with greater severity of pulmonary deficits observed in chest imaging. Breath-holding physiology may

have utility for rapid identification of COVID-19 patients at elevated risk of respiratory failure particularly in circumstances where resources are limited.

AUTHOR CONTRIBUTION
Ludovico Messineo: Conceptualization (equal); formal analysis (lead); investigation (lead); methodology (lead); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (lead); writing – review and editing (equal). **Francesco Fanfulla:** Data curation (equal); investigation (equal); project administration (equal); supervision (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Leonardo Pedroni:** Data curation (equal); software (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Floriana Pini:** Data curation (equal); investigation (equal); visualization (equal); writing – review and editing (equal). **Andrea Borghesi:** Formal analysis (equal); investigation (equal); methodology (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Salvatore Golemi:** Formal analysis (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Guido Vailati:** Data curation (equal); visualization (equal); writing – review and editing (equal). **Kayla Kerlin:** Methodology (supporting); supervision (supporting); validation (supporting); visualization (equal); writing – review and editing (equal). **Atul Malhotra:** Conceptualization (supporting); supervision (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Luciano Corda:** Data curation (equal); investigation (equal); supervision (equal); validation (equal); visualization (equal); writing – review and editing (equal). **Scott Sands:** Conceptualization (equal); formal analysis (equal); investigation (equal); methodology (equal); software (equal); supervision (equal); validation (equal); visualization (equal); writing – original draft (supporting); writing – review and editing (equal).

CONFLICTS OF INTEREST
Francesco Fanfulla reports honoraria fee from Jazz Medical and GSK for scientific lectures. Atul Malhotra reports income related to medical education from Equillium, Corvus, Jazz and Livanova; ResMed provided a philanthropic donation to UC San Diego. Scott Sands reports grants and personal fees from Apnimed, personal fees from Nox Medical, personal fees from Merck, personal fees from Inspire, grants from Prosomnus, grants from Dynaflex and outside the submitted work. In addition, Scott Sands has a patent pending related to wearable oximetry technology for sleep apnoea diagnosis/phenotyping, unrelated to breath-holding.

DATA AVAILABILITY STATEMENT
All the individual participant and summary data collected during the trial will be shared upon request, after de-identification. Additionally, study protocol, statistical analysis plan and informed consent will be made available. Data will be available immediately following publication and ending 5 years following article publication with researchers who provide a methodologically sound proposal to achieve aims in such approved proposal. Proposals should be directed to Ludovico Messineo ludovico.messineo@yahoo.it; access will require a data use agreement.

HUMAN ETHICS APPROVAL DECLARATION
This study was approved by the Italian Scientific Institutes Maugeri Ethics Committee (number: 2505) on 1 December 2020, which was the reference centre. The study was also approved by Brescia local Ethics Committee (number: 4573) on 12 January 2020 and Chiari local Ethics Committee (number: 4790) on 18 May 2020. All participants provided verbal or written informed consent upon enrolment. The study conformed to the standards set by the latest revision of the Declaration of Helsinki.

ORCID
Ludovico Messineo https://orcid.org/0000-0002-4142-9124

REFERENCES
1. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019—COVID-NET, 14 states, March 1-30, 2020. MMWR Morb Mortal Wkly Rep. 2020; 69(15):458–64.
2. Leon TM, Dorabawila V, Nelson L, Lutterloh E, Bauer UE, Backenson B, et al. COVID-19 cases and hospitalizations by COVID-19 vaccination status and previous COVID-19 diagnosis—California and New York, May-November 2021. MMWR Morb Mortal Wkly Rep. 2022;71(4):125–31.
3. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054–62.
4. Bosch W, Cowart JB, Bhakta S, Carter RE, Wadei HM, Shah SZ, et al. COVID-19 vaccine-breakthrough infections requiring hospitalization in Mayo Clinic Florida through August 2021. Clin Infect Dis. 2021;ciab932. https://doi.org/10.1093/cid/ciab932
5. Brosh-Nissimov T, Orenbuch-Harroch E, Chowers M, Elbaz M, Nesher L, Stein M, et al. BNT162b2 vaccine breakthrough: clinical characteristics of 152 fully vaccinated hospitalized COVID-19 patients in Israel. Clin Microbiol Infect. 2021;27:1652–7.
6. Brown CM, Vostok J, Johnson H, Burns M, Gharpure R, Sami S, et al. Outbreak of SARS-CoV-2 infections, including COVID-19 vaccine breakthrough infections, associated with large public gatherings—Barnstable County, Massachusetts, July 2021. MMWR Morb Mortal Wkly Rep. 2021;70(31):1059–62.
7. Liu C, Lee J, Ta C, Soroush A, Rogers JR, Kim JH, et al. A retrospective analysis of COVID-19 mRNA vaccine breakthrough infections—risk factors and vaccine effectiveness. medRxiv. 2021.

8. Team CC-VBCI. COVID-19 vaccine breakthrough infections reported to CDC—United States, January 1-April 30, 2021. MMWR Morb Mortal Wkly Rep. 2021;70(21):792–3.

9. Tartof SY, Slezak JM, Fischer H, Hong V, Ackerson BK, Ranasinghe ON, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet. 2021;398(10309):1407–16.

10. Cohn BA, Cirillo PM, Murphy CC, Krigbaum NY, Wallace AW. SARS-CoV-2 vaccine protection and deaths among US veterans during 2021. Science. 2021;375:331–6.

11. Callaway E. Heavily mutated omicron variant puts scientists on alert. Nature. 2021;600(7887):21.

12. Volz E, Mishra S, Chand M, Barrett JC, Johnson R, Geidelberg L, et al. Assessing transmissibility of SARS-CoV-2 lineage B.1.1.7 in England. Nature. 2021;593(7858):266–9.

13. McCay L. Omicron: urgent action needed on NHS staffing crisis. BMJ. 2022;376:o18.

14. Kadri SS, Simpson SQ. Potential implications of SARS-CoV-2 delta variant surges for rural areas and hospitals. JAMA. 2021;326(11):1003–4.

15. Rosenbaum L. Facing Covid-19 in Italy—ethics, logistics, and therapeutics on the epidemic's front line. N Engl J Med. 2020;382(20): 1873–5.

16. Messineo L, Taranto-Montemurro L, Azarbarzin A, Oliveira Marques MD, Calianese N, White DP, et al. Breath-holding as a means to estimate the loop gain contribution to obstructive sleep apnoea. J Physiol. 2018;596(17):4043–56.

17. Messineo L, Perger E, Corda L, Joosten SA, Fanfulla F, Pedroni L, et al. Breath-holding as a novel approach to risk stratification in COVID-19. Crit Care. 2021;25(1):208.

18. Tobin MJ, Laghi F, Jubran A. Why COVID-19 silent hypoxemia is baffling to physicians. Am J Respir Crit Care Med. 2020;202(3): 356–60.

19. Wynants L, Van Calster B, Collins GS, Riley RD, Heinze G, Schuit E, et al. Prediction models for diagnosis and prognosis of covid-19: systematic review and critical appraisal. BMJ. 2020;369:m1328.

20. Knight SR, Ho A, Pius R, Buchan I, Carson G, Drake TM, et al. Risk stratification of patients admitted to hospital with covid-19 using the ISARIC WHO Clinical Characterisation Protocol: development and validation of the 4C Mortality Score. BMJ. 2020;370: m3339.

21. Duca A, Piva S, Foca E, Latronico N, Rizzi M. Calculated decisions: Brescia-COVID respiratory severity scale (BCRSS)/algorithm. Emerg Med Pract. 2020;22(Suppl 5):CD1–2.

22. WHO. Rational use of personal protective equipment for coronavirus disease (COVID-19) and considerations during severe shortages 2020. https://www.who.int/publications/i/item/rational-use-of-personal-protective-equipment-for-coronavirus-disease-(covid-19)-and-considerations-during-severe-shortages.

23. Borghesi A, Maroldi R. COVID-19 outbreak in Italy: experimental chest X-ray scoring system for quantifying and monitoring disease progression. Radiol Med. 2020;125(5):509–13.

24. Colombi D, Bodini FC, Petrini M, Maffi G, Morelli N, Milanese G, et al. Well-aerated lung on admitting chest CT to predict adverse outcome in COVID-19 pneumonia. Radiology. 2020;296(2): E86–96.

25. Sverzellati N, Ryerson CJ, Milanese G, Renzoni EA, Volpi A, Spagnolo P, et al. Chest radiography or computed tomography for COVID-19 pneumonia? Comparative study in a simulated triage setting. Eur Respir J. 2021;58(3):2004188.

26. Sands SA, Edwards BA, Kelly VJ, Davidson MR, Wilkinson MH, Berger PJ. A model investigation of the impact of ventilation-perfusion mismatch on oxygenation during apnea in preterm infants. J Theor Biol. 2010;264:657–62. https://doi.org/10.1016/j.jtbi.2010.03.041

27. Findley LJ, Ries AL, Tisi GM, Wagner PD. Hypoxemia during apnea in normal subjects: mechanisms and impact of lung volume. J Appl Physiol. 1983;55(6):1777–83.

28. Sands SA, Edwards BA, Kelly VJ, Davidson MR, Wilkinson MH, Berger PJ. A model analysis of arterial oxygen desaturation during apnea in preterm infants. PLoS Comput Biol. 2009;5(12):e1000588.

29. Mantha S, Tripuraneni SL, Roizen MF, Fleisher LA. Proposed modifications in the 6-minute walk test for potential application in patients with mild COVID-19: a step to optimize triage guidelines. Anesth Analg. 2020;131(2):398–402.

30. Pandit R, Vaity C, Mulakavalupil B, Matthew A, Sabnis K, Joshi S. Unmasking hypoxia in COVID 19—six minute walk test. J Assoc Physicians India. 2020;68(9):50–1.

31. Kalin A, Javid B, Knight M, Inada-Kim M, Greenhalgh T. Direct and indirect evidence of efficacy and safety of rapid exercise tests for exertional desaturation in Covid-19: a rapid systematic review. Syst Rev. 2021;10(1):77.

32. Goodacre S, Thomas B, Lee E, Sutton L, Loban A, Waterhouse S, et al. Post-exertion oxygen saturation as a prognostic factor for adverse outcome in patients attending the emergency department with suspected COVID-19: a substudy of the PRIEST observational cohort study. Emerg Med J. 2021;38(2):88–93.

33. Villadiego J, Ramírez-Lorca R, Cala F, Labandeira-García JL, Esteban M, Toledo-Aral JJ, et al. Is carotid body infection responsible for silent hypoxemia in COVID-19 patients? Function. 2020;2(1): zqaa032.

34. Schagatay E, van Kampen M, Emanuelsson S, Holm B. Effects of physical and apnea training on apneic time and the diving response in humans. Eur J Appl Physiol. 2000;82(3):161–9.

35. Schiaffino S, Albano D, Cozzi A, Messina C, Arioli R, Bna C, et al. CT-derived chest muscle metrics for outcome prediction in patients with COVID-19. Radiology. 2021;300(2):E328–E36.

36. Kalra MK, Homayounieh F, Arru C, Holmberg O, Vassileva J. Chest CT practice and protocols for COVID-19 from radiation dose management perspective. Eur Radiol. 2020;30(12):6554–60.

37. Brun AL, Gence-Breney A, Trichereau J, Ballester MC, Vasse M, Chabi ML, et al. COVID-19 pneumonia: high diagnostic accuracy of chest CT in patients with intermediate clinical probability. Eur Radiol. 2021;31(4):1969–77.

38. Agostini A, Floridi C, Borgheresi A, Badaloni M, Esposto Pirani P, Terilli F, et al. Proposal of a low-dose, long-pitch, dual-source chest CT protocol on third-generation dual-source CT using a tin filter for spectral shaping at 100 kVp for CoronaVirus disease 2019 (COVID-19) patients: a feasibility study. Radiol Med. 2020;125(4):365–73.

39. Malas MB, Naazie IN, Elsayed N, Mathlouthi A, Marmor R, Clary B. Thromboembolism risk of COVID-19 is high and associated with a higher risk of mortality: a systematic review and meta-analysis. EClinicalMedicine. 2020;29:100639.

40. Wang S, Zha Y, Li W, Wu Q, Li X, Niu M, et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. Eur Respir J. 2020;56(2):2000775.

41. Angeli E, Dalto S, Marchese S, Setti L, Bonacina M, Galli F, et al. Prognostic value of CT integrated with clinical and laboratory data during the first peak of the COVID-19 pandemic in northern Italy: a nomogram to predict unfavorable outcome. Eur J Radiol. 2021;137: 109612.

42. Bernheim A, Mei X, Huang M, Yang Y, Fayad ZA, Zhang N, et al. Chest CT findings in coronavirus Disease-19 (COVID-19): relationship to duration of infection. Radiology. 2020;295(3):200463.

43. Wong HYF, Lam HYS, Fong AH, Leung ST, Chin TW, Lo CSY, et al. Frequency and distribution of chest radiographic findings in patients positive for COVID-19. Radiology. 2020;296(2):E72–8.

44. Xie J, Hungerford D, Chen H, Abrams ST, Li S, Wang G, et al. Development and external validation of a prognostic multivariable model on admission for hospitalized patients with COVID-19. medRxiv. 2020:2020.03.28.20045997.

45. Zhang H, Shi T, Wu X, Zhang X, Wang K, Bean D, et al. Risk prediction for poor outcome and death in hospital in-patients with COVID-19: derivation in Wuhan, China and external validation in London, UK 2020:2020.04.28.20082222.

46. Sanità ISd. Prevalenza e distribuzione delle varianti di SARs-CoV-2 di interesse per la sanità pubblica in Italia 2021. https://www.iss.it/documents/20126/0/Blettino+varianti+n.+9_17+settembre+2021.pdf/484b7aa2-2c0c-b109-4c31-087ed5c7b5af?t=1631890444760.

47. Iuliano AD, Brunkard JM, Boehmer TK, Peterson E, Adjei S, Binder AM, et al. Trends in disease severity and health care utilization during the early omicron variant period compared with previous SARS-CoV-2 high transmission periods—United States, December 2020-January 2022. MMWR Morb Mortal Wkly Rep. 2022;71(4): 146–52.