Differences in acoustic features of cough by pneumonia severity in patients with COVID-19: a cross-sectional study

Clare Davidson, Osw Antonio Caguana, Manuel Lozano-García, Mariela Arita Guevara, Luis Estrada-Petrocelli, Ignasi Ferrer-Lluis, Yolanda Castillo-Escario, Pilar Ausín, Joaquim Gea, Raimon Jané

Differences in acoustic features of cough by pneumonia severity in patients with COVID-19: a cross-sectional study

Clare Davidson[1*], Osw Antonio Caguana[2*], Manuel Lozano-García[3,1,4], Mariela Arita Guevara[2], Luis Estrada-Petrocelli[1,4,5], Ignasi Ferrer-Lluis[1,3,4], Yolanda Castillo-Escario[1,3,4], Pilar Ausín[2], Joaquim Gea[2ǂ], Raimon Jané[3,1,4ǂ]

(*) These authors share the first author position since their contribution was equivalent.

(ǂ) These two authors share the senior author position.

1. Institute for Bioengineering of Catalonia (IBEC), The Barcelona Institute of Science and Technology (BIST)., Barcelona, Spain.

2. Respiratory Medicine Dept., Hospital del Mar – IMIM. MELIS, Universitat Pompeu Fabra. CIBERES, ISCiii. BRN. Barcelona, Spain.

3. Universitat Politècnica de Catalunya (UPC)—Barcelona Tech, Barcelona, Spain.

4. Centro de Investigación Biomédica en Red en Bioingeniería, Biomateriales y Nanomedicina (CIBER-BBN), Barcelona, Spain.

5. Universidad Latina de Panamá, Facultad de Ingeniería, Panama, Panama.

Corresponding author:

Clare Davidson, cdavidson@ibecbarcelona.eu, +34 722 21 27 73

Summary to publicise via social media:

Acoustic analysis of cough sounds recorded via smartphone in COVID-19 patients reveals features of cough that could potentially be used to provide a fast, easy, cost-effective way to identify patients' disease severity at home or in any healthcare setting.

Abstract:

Background: Acute respiratory syndrome due to coronavirus 2 (SARS-CoV-2) is characterised by heterogeneous levels of disease severity. It is not necessarily apparent whether a patient will develop a severe disease or not. This cross-sectional study explores whether acoustic properties of the cough sound of patients with coronavirus disease (COVID-19), the illness caused by SARS-CoV-2, correlate with their disease and pneumonia severity, with the aim of identifying patients with a severe disease.

Methods: Voluntary cough sounds were recorded using a smartphone in 70 COVID-19 patients within the first 24 hours of their hospital arrival, between April 2020 and May 2021. Based on gas exchange abnormalities, patients were classified as mild, moderate, or severe. Time- and frequency-based variables were obtained from each cough effort and analysed using a linear mixed-effects modelling approach.

Results: Records from 62 patients (37% female) were eligible for inclusion in the analysis, with mild, moderate, and severe groups consisting of 31, 14 and 17 patients respectively. 5 of the parameters examined were found to be significantly different in the cough of patients at different disease levels of severity, with a further 2 parameters found to be affected differently by the disease severity in men and women.

Conclusions: We suggest that all these differences reflect the progressive pathophysiological alterations occurring in the respiratory system of COVID-19 patients, and potentially would provide an easy and cost-effective way to initially stratify patients, identifying those with more severe disease, and thereby most effectively allocate healthcare resources.

Introduction

The global coronavirus disease 2019 (COVID-19) pandemic caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [1] continues to be a major health problem. Although most people affected by COVID-19 now have mild-to-moderate symptoms and recover within a few

weeks, some that develop more severe disease and pneumonia often have a poorer prognosis. It is not known with certainty which factors predispose to severe disease, although certain genetic variants have been implicated [2]. Other risk factors including age, comorbidities, and determinants of cardiovascular risk, have also been identified [3]. There is also evidence that males are more susceptible to severe disease and death from COVID-19 [4, 5]. Although it has been suggested that the immune and inflammatory response may contribute to this gender disparity, the underlying pathophysiological mechanisms have not been fully elucidated.

Much research time and money has been invested into finding ways to obtain an early COVID-19 diagnosis. In this regard, laboratory methods, and the use of imaging techniques [6], [7], statistical models [8] and artificial intelligence [9] have been investigated. The currently accepted gold-standard diagnostic test – reverse transcription polymerase chain reaction (RT-PCR) – is widely available and relatively accessible [10]. However, although risk stratification protocols have been developed [11], potential diagnostic and prognostic tools are mostly based on relatively expensive and in many scenarios, difficult-to-access imaging methods (radiography, ultrasound, computed tomography (CT)) [12, 13]. There is a clinical need for a simpler and more widely available prognostic tool that would enable healthcare providers to identify patients who have developed or are at risk of developing a severe disease, thereby facilitating triaging of patients and early intervention even at a patient's home or primary care centres [14–16].

The analysis and interpretation of cough sounds in the initial stages of COVID-19 disease could potentially provide a predictive tool that would meet these criteria. A dry cough is one of the most common symptoms of COVID-19, which occurs during the initial disease phase in up to 70% of patients [17, 18]. To date, several studies have applied machine learning paradigms to the acoustic properties of cough sounds to develop a screening or diagnostic tool for COVID-19 [19–22]. Many of these studies have leveraged the recording capabilities of the ubiquitous and easy-to-use smartphone to collect data, often via crowdsourcing techniques [9]. Such devices are available to a large proportion of the population and allow cost-effective recording of coughs using built-in microphones, even outside more sophisticated health care settings [23].

We hypothesised that the acoustic properties of the cough sound of patients with COVID-19 would differ with disease and pneumonia severity. Given the different effects of COVID-19 on males and females, we expected that the results would differ by gender. In addition, we postulated that several other variables: age, smoking status, pre-existing respiratory conditions, length of time with symptoms and fraction of inspired oxygen (FiO_2) could also potentially affect this relationship. To

test this hypothesis, here we explore the correlation between the frequency content of cough sounds recorded via smartphone with the disease and pneumonia severity in COVID-19 patients.

Material and methods:

Voluntary cough sounds were recorded in 70 COVID-19 patients over the age of 18 years, whose disease symptoms had been present for 15 days or less, within the first 24 hours (in some exceptional cases 48 hours) of their arrival to our teaching hospital. We used a cross-sectional study design, with data collection taking place on 21 different dates between April 2020 and May 2021. The sample size was determined by the number of participants available on these dates. The participants were divided into three groups according to the severity of their disease at the time of cough recording. The *mild group* consisted of patients without pneumonia, the *moderate group* were patients with pneumonia but not requiring supplemental oxygen, and *severe patients* showed moderate or severe pneumonia that necessitated oxygen therapy with invasive or non-invasive respiratory support [24]. The pre-existing respiratory conditions that were present in some patients were asthma, chronic obstructive pulmonary disease, and interstitial pulmonary fibrosis. The respiratory rate of the patients was between 18 and 24 breaths per minute, and there were no apparent limitations to their production of voluntary cough sounds. The study was conducted in accordance with the Declaration of Helsinki and approved by the institutional Ethics Committee (CEIM, ref. 10231l). Fully informed written consent was obtained from patients or their relatives prior to their inclusion in the study. In some cases, verbal consent was initially obtained, as recommended by the Ethics Committee, and was further completed with the written consent.

Voluntary cough sounds were recorded by respiratory medicine specialists with a smartphone (Samsung Galaxy S21). Recordings were made in a room with as little background noise as possible. Patients were instructed to take a deep breath and then cough voluntarily; this often caused them to trigger an involuntary cough. Patients coughed 3-4 times in the direction of the smartphone, which was positioned 15-20 cm from their mouth. Patients who required a low-flow oxygen device were asked to remove the mask briefly to perform the manoeuvres. Cough sounds were acquired and sampled at 48kHz using the built-in hardware on the smartphone and the Easy Voice Recorder application (available at Google Play Store). For infection control purposes, the smartphone was encased in a disposable latex cover prior to each recording. We previously confirmed that the addition of this cover did not affect the fidelity of the cough recording by comparing the temporal, spectral, and time-frequency characteristics of audio signals recorded with and without the cover.

In this study, a single expiratory effort has been labelled as a 'cough effort' (CE), with a 'cough bout' consisting of two or more CEs following a single initial inspiration [25]. This is illustrated in Figure 1(a). Each CE can be further segmented into three constituent parts: a first sound (CS_1), an intermediate part (C_{INT}), and if present, a second sound (CS_2). This segmentation was performed manually using both the aural and visual representation of each cough sound and is illustrated in Figure 1(b) [26]. The cough sound occurs during the expulsive phase of the cough, with the first sound happening at the moment of glottal opening. The intermediate part follows this and represents the steady-state flow of air with the glottis open. The second cough sound, which is not always present, occurs at the end of the expulsive phase as the glottis narrows. Cough recordings were included as valid in the database for further analysis only if a minimum of one valid cough bout, consisting of a minimum of two CEs, could be identified and isolated from the recording. Single cough efforts were excluded as the results of a single effort may be random and thus not representative of the cough sounds of a participant, and to ensure a balanced data set (equal number of coughs in first and second position in the bout).

Individual CEs were identified in each recording by visual and aural inspection of the recorded signals [27]. A flowchart outlining the analysis methodology is shown in Figure 2. A total of 459 CEs were isolated, with a median of 6 CEs (range 5 – 10) from each recording. Information on the CE number within the recording as well as the CE position within a bout were also annotated. The CEs were filtered using a 20th order Chebyshev Type II low-pass filter with a cut-off frequency of 6 kHz to minimise background noise. The envelope of each identified CE was then calculated using the root mean square. This was used as an aid to identify and manually split each CE into its above-mentioned constituent parts – CS_1, C_{INT}, and CS_2 when this was present. For each CE, as well as its constituent parts, the power spectrum was estimated using Welch's method with a Hanning window and a 50% overlap applied to compute the modified periodograms. Several time- and frequency-based parameters of each whole CE signal and its constituent parts were obtained and analysed (Table 1). Data analysis was performed offline using custom developed scripts in Matlab [28].

Statistical analysis was performed in RStudio [29]. The relationship between each parameter and disease severity was investigated with a linear mixed-effects model with maximum likelihood optimisation using the *lme4* library [30]. The mean value of each parameter for each CE position (nesting level one) was nested within each subject (nesting level two), which was in turn nested according to disease severity (nesting level three). Disease severity, gender and CE position were entered as fixed effects into the model. An interaction term (disease severity * gender) was included in the model to investigate whether the parameter being examined was affected differently by disease severity in males and females. An intercept for individual subjects was included as a random

effect to account for differences between subjects. Several further models were then defined, each with an additional fixed effect added to the main model. These were the patients' age, smoking status, length of time with symptoms, presence/absence of a pre-existing respiratory condition, and FiO_2. In each case, visual inspection of residual plots did not reveal any obvious deviations from homoscedasticity or normality. The variance inflation factor (VIF) was calculated for each independent variable to ensure that there was no collinearity between them (VIF < 5). A type II ANOVA with F-tests and p-values using Satterthwaite's method for denominator degrees-of-freedom and F-statistic was applied to test whether the parameters had a statistically significant effect on the derived model. Non-parametric Kruskal-Wallis and Wilcoxon Rank Sum tests were applied as post-hoc tests, with Benjamini and Hochberg p-value adjustment. In those cases where a significant interaction effect between disease severity and gender was observed, male and female data were also examined separately. An α value of 0.05 was used to indicate significance throughout.

Results

A total of 70 participant recordings were initially examined for eligibility. Of these, recordings from 6 individuals were excluded due to technical problems with the recording quality (5 recordings had (inadvertently) been recorded at too low sampling frequency and a sixth was excluded due to the presence of a second person coughing simultaneously), and a further 2 recordings were also excluded as no clearly discernible cough sounds were present. The remaining 62 recordings were deemed eligible and included in all analyses. The main participant characteristics and relevant clinical data is presented in Table 2.

All participants had a minimum of one CE at position one (CE_1) and one at position two (CE_2). Only 33 participants (53%) had a CE at position three (CE_3). Therefore, only CE_1s and CE_2s were included in the analysis, to ensure balanced representation of each participant, and to allow for the effect of CE position within the bout on the features extracted to be examined.

As for the effect of gender, higher frequency content was found in female coughs than in male coughs. Cough sounds in general mirror the natural expected frequency content of the voice of males and females, with overall frequency content of female coughs higher than that of male coughs.

Five of the parameters examined were found to be significantly different in the cough recordings of patients at different disease levels (Figure 3 & Figure 4). In the whole CE signal, these parameters were the frequency variability (F_{VAR}) (760.3 vs 767.4 vs 614.9 Hz, p = 0.0031) and peak frequency (F_{PK})

(473.9 vs 340.1 vs 610.3 Hz, p = 0.0025). Values given within parentheses are the median values of the mild, moderate, and severe groups, respectively, with this order followed in all the results here. In CS_1, the F_{VAR} (729.5 vs 601.9 vs 526.7 Hz, p = 0.0010) and frequency of maximum energy (F_{MAX}) (2191.8 vs 1898.4 vs 1620.4 Hz, p = 0.0130) differed significantly, and in C_{INT}, the interquartile range (F_{IQR}) (674.7 vs 825.7 vs 527.6 Hz, p = 0.0260) also differed significantly. Pairwise comparisons using the Wilcoxon Rank Sum test with continuity correction revealed significant differences for all five parameters between individuals with mild and severe disease. A significant difference for F_{VAR} and F_{PK} in the whole CE signal, and the F_{IQR} of C_{INT} between those with moderate and severe disease was also found.

A significant interaction term for disease severity and gender was found for two parameters: F_{VAR} and F_{MAX} of C_{INT} (Figure 5). Moreover, significant differences were observed in F_{VAR} of C_{INT} for males (697.2 vs 752.0 vs 588.6 Hz, p = 0.0131) and females (1054.5 vs 768.8 vs 597.5 Hz, p < 0.0001), and in F_{MAX} of C_{INT} for females only (3567.0 vs 2691.0 vs 2190.0 Hz, p < 0.001). Pairwise comparisons using the Wilcoxon Rank Sum test with continuity correction revealed significant differences for F_{VAR} of C_{INT} between female patients at all disease levels, and for F_{MAX} of C_{INT} between female patients with mild and moderate, and mild and severe disease. For male patients a significant difference was observed for F_{VAR} in C_{INT} between individuals with moderate and severe disease.

The position of the CE within the cough bout had a significant effect on the duration of the whole CE signal, and the duration of CS_1 and C_{INT} individually. However, there was no significant difference found between the frequency parameters reported here for CE positions one and two. The addition of a fixed effect of either patients' age, smoking status, length of time with symptoms, presence/absence of a pre-existing respiratory condition, or FiO_2, to the main model, was found to have no significant effect on the model at a significance level of α = 0.05.

Discussion

This study describes the relationship between frequency-based features of the cough sound in COVID-19 and the disease and pneumonia severity in the patient. These relationships were explored, considering patient and disease profiles (gender, age, and smoking status, as well as duration of COVID-19 symptoms, presence or absence of a pre-existing respiratory condition and the oxygen requirements) using linear mixed-effects models. The analysis of cough recordings is a relatively easy way to get information about some diseases in the respiratory system. A qualitative assessment of cough sounds may be done by a medical professional in usual care scenarios. However, this relatively coarse assessment is subjective and depends on the expertise and hearing acuity of the

professional involved. Healthcare professionals usually just differentiate dry vs. productive cough; in fact, this is the most common comment in standard clinical records. Therefore, although a high-level distinction between disease types may be observed, more subtle nuances of cough sounds may be missed, and healthcare professionals may encounter difficulties in diagnosing from cough sounds [31]. Automatic algorithms can help to extract objective information from cough sounds and thus simplify the process and support the medical staff.

In our quantitative analysis, five frequency-based features (F_{VAR} and F_{PK} of the whole CE signal, F_{VAR} and F_{MAX} of CS_1 and F_{IQR} of C_{INT}), were found to differ significantly with disease severity, the classification of which is based on the presence and/or severity of pneumonia [24]. We suggest that these differences reflect the progressive pathophysiological alterations of the respiratory system in patients with COVID-19 [32]. Differences have been previously noted in chest CT scans between patients with mild and severe/critical disease [13]. Although similar analysis of acoustic properties of cough sounds has been used to diagnose respiratory illnesses [33–35], we are not aware of any studies that explore a possible relationship between cough sounds and varying disease severity levels of a respiratory illness.

Two further frequency-based features, F_{VAR} and F_{MAX} of C_{INT}, were observed to be affected differently in male and female patients by disease severity. C_{INT} occurs between CS_1 and CS_2 and is the part of the cough sound that is produced via steady-state airflow with the glottis open. It is possible that the pathophysiology of COVID-19 differs between both genders in this part due to well-known differences of male and female anatomy, which would be reflected in the sound differences we observed.

The data used in the present study consists of a clinically recorded and validated dataset, collected from a relatively large cohort of well characterised patients. The cough recordings were acquired with an easy-to-use smartphone application in an early period of a patients first contact with the health system, and by the healthcare professional caring for the patient. This helped to ensure that the recordings were of a consistent high quality across participants. The availability of relevant patient information enabled us to explore the effect of possible covariates – age, smoking status, length of time with symptoms, presence/absence of a pre-existing respiratory condition, and required FiO_2 on the results obtained in our analysis. These strengths offer distinct advantages over other studies that use datasets that have been crowdsourced or collected using less stringent methodology. In addition, the use of the smartphone enables the cough recordings to be acquired in virtually any setting, thus overcoming limitations posed by location-dependent imaging, and other, techniques.

There are also some possible limitations to our study. Spontaneous cough recording could be considered the optimal way to predict the pathophysiological situation in a respiratory patient. However, this can be difficult to acquire, as patients can have long periods without this spontaneous effort occurring. Therefore, we collected voluntary, induced coughs, which are easy to perform and have previously been validated as a good surrogate measure of the spontaneous cough from an acoustic perspective [36]. Our study included the analysis of CEs from positions one (CE_1) and two (CE_2) within a cough bout. The definition of a classical cough includes an inhalation prior to the cough sound occurring [37]. Therefore, the second cough sound in a bout is likely to be an expiration reflex (ER) rather than a true cough. However, although the distinction exists, as the two sounds are indistinguishable to the human ear, for clinical purposes no distinction was made between them. Our results suggest that the frequency content of the classical cough (CE_1) and the ER (CE_2) does not differ, but we noted that the duration of the ER appears shorter than that of the classical cough. Finally, although we found some apparent differences in the acoustic features of male and female cough sounds, our database was not completely balanced, consisting of 37% female patients.

Our study highlights acoustic features of the cough sound in COVID-19 patients that differ significantly with disease and pneumonia severity. The results obtained suggest that it might be possible to identify and predict the severity and extent of COVID-19 from the cough sound of a particular patient. However, despite the significant differences reported, it must be noted that there is a substantial variability and overlap between parameters from patients with different COVID-19 severity. Moreover, these parameters can also vary within individual patients, and this might be another important source of variability. For example, the mean of the intra-subject standard deviation for the F_{VAR} of the whole CE signal was 106 Hz in the mild group, 87 Hz in the moderate group, and 75 Hz for the severe group. Interestingly, the intra-subject variation was lower in the severe group than in the mild group. The potential of the proposed features for classification purposes has not been studied yet but remains a topic for further research. Using machine learning techniques and perhaps adding some extra features, the potential of this approach for discriminating different severities could be confirmed. This could result in an early stratification and prediction of probable clinical outcomes to triage correctly and allocate healthcare resources accordingly, which would be of huge benefit for both patient and healthcare providers. Further studies would elucidate if this methodology may also be extendable to long-COVID to analyse if the evolution of the cough signal can reflect the presence or severity of respiratory sequelae (organising pneumonia, interstitial fibrosis, hyperreactivity).

TABLE 1: TIME AND FREQUENCY PARAMETERS OBTAINED FROM THE COUGH RECORDINGS OF STUDY PARTICIPANTS.

Parameter	Description
Dur	Duration of segment in time
F_{MED}	Median frequency
F_{MID}	Mean spectral frequency
F_{PK}	Frequency at point of maximum spectral energy
F_{VAR}	Frequency variability (std. dev)
F_{Q1}	First quartile frequency
F_{Q3}	Third quartile frequency
F_{IQR}	Interquartile range of frequency ($F_{Q3} - F_{Q1}$)
F_{MAX}	Frequency at 95% spectral energy
Skew	Skewness
Kurt	Kurtosis

TABLE 2: ANTHROPOMETRIC AND RELEVANT CLINICAL DATA OF STUDY PARTICIPANTS.

Disease Severity*	Subjects (n)	Female n (%)	Age (years)	Smoker† n (%)	Pre-existing respiratory condition‡ n (%)	Time with Symptoms (days)
Mild	31	11 (35)	52.0 [44.0 – 60.5]	6 (19)	3 (10)	9.0 [7.0 – 10.5]
Moderate	14	5 (36)	51.5 [46.0 – 60.0]	3 (21)	2 (14)	9.5 [6.3 – 11.8]
Severe	17	7 (41)	54.0 [43.0 – 64.0]	2 (12)	2 (12)	8.0 [6.0 – 10.0]

Data is presented as n (%) or median [interquartile range] unless otherwise stated.
* **Mild** = no pneumonia, **Moderate** = pneumonia (not requiring supplemental oxygen), **Severe** = moderate/severe pneumonia (requiring supplemental oxygen)
† The smoker category includes both current and ex-smokers.
‡ Pre-existing respiratory conditions included here are asthma, COPD and Interstitial pulmonary fibrosis (IPF).

Bibliography

1.   Zhu N, Zhang D, Wang W, et al. A novel coronavirus from patients with pneumonia in China, 2019. New England journal of medicine 2020;382:727–733.

2.   Kousathanas A, Pairo-Castineira E, Rawlik K, et al. Whole genome sequencing reveals host factors underlying critical Covid-19. Nature 2022;https://doi.org/10.1038/s41586-022-04576-6.

3.   Du RH, Liang LR, Yang CQ, et al. Predictors of mortality for patients with COVID-19 pneumonia caused by SARS-CoV-2: a prospective cohort study. European Respiratory Journal 2020;55.

4.   Takahashi T, Ellingson MK, Wong P, et al. Sex differences in immune responses that underlie COVID-19 disease outcomes. Nature 2020;588:315.

5.   Williamson EJ, Walker AJ, Bhaskaran K, et al. Factors associated with COVID-19-related death using OpenSAFELY. Nature 2020;584:430.

6.   Kanne JP, Bai H, Bernheim A, et al. COVID-19 imaging: what we know now and what remains unknown. Radiology 2021;299:E262–E279.

7.   Dong D, Tang Z, Wang S, et al. The role of imaging in the detection and management of COVID-19: a review. IEEE Rev Biomed Eng 2020;14:16–29.

8.   Jehi L, Ji X, Milinovich A, et al. Individualizing Risk Prediction for Positive Coronavirus Disease 2019 Testing: Results From 11,672 Patients. Chest 2020;158:1364–1375.

9.   Laguarta J, Hueto F, Subirana B. COVID-19 Artificial Intelligence Diagnosis Using only Cough Recordings. IEEE Open J Eng Med Biol 2020;1:275–281.

10. Corman VM, Landt O, Kaiser M, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. Eurosurveillance 2020;25.

11. Knight SR, Ho A, Pius R, 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;370:22.

12. Zhang K, Liu X, Shen J, et al. Clinically Applicable AI System for Accurate Diagnosis, Quantitative Measurements, and Prognosis of COVID-19 Pneumonia Using Computed Tomography. Cell 2020;181:1423–1433.

13. Li K, Wu J, Wu F, et al. The Clinical and Chest CT Features Associated with Severe and Critical COVID-19 Pneumonia. Invest Radiol 2020;55:327–331.

14. Goyal D, Inada-Kim M, Mansab F, et al. Improving the early identification of COVID-19 pneumonia: a narrative review. BMJ Open Respir Res 2021;8:e000911.

15. Feng Z, Yu Q, Yao S, et al. Early prediction of disease progression in COVID-19 pneumonia patients with chest CT and clinical characteristics. Nat Commun 2020;11:1–9.

16. Wynants L, van Calster B, Collins GS, et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. BMJ 2020;369:m1328.

17. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. The Lancet 2020;395:1054–1062.

18. Guan W, Ni Z, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. New England journal of medicine 2020;382:1708–1720.

19. Pahar M, Klopper M, Warren R, Niesler T. COVID-19 cough classification using machine learning and global smartphone recordings. Comput Biol Med 2021;135:104572.

20. Imran A, Posokhova I, Qureshi HN, et al. AI4COVID-19: AI enabled preliminary diagnosis for COVID-19 from cough samples via an app. Inform Med Unlocked 2020;20:100378.

21. Mouawad P, Dubnov T, Dubnov S. Robust detection of COVID-19 in cough sounds. SN Comput Sci 2021;2:1–13.

22. Han J, Xia T, Spathis D, et al. Sounds of COVID-19: exploring realistic performance of audio-based digital testing. NPJ Digit Med 2022;5:1–9.

23. Hall JI, Lozano M, Estrada-Petrocelli L, Birring S, Turner R. The present and future of cough counting tools. J Thorac Dis 2020;12:5207–5223.

24. World Health Organization. Living guidance for clinical management of COVID-19: living guidance, 23 November 2021. World Health Organization, 2021.

25. Morice AH, Fontana GA, Belvisi MG, et al. ERS guidelines on the assessment of cough. European Respiratory Journal 2007;29:1256–1276.

26. Korpáš J, Sadloňová J, Vrabec M. Analysis of the cough sound: An overview. Pulm Pharmacol 1996;9:261–268.

27. Audacity Team (2021). Audacity(R): Free Audio Editor and Recorder (Computer Application).

28. MATLAB. version 9.9.0 (R2020b). 2020. The MathWorks Inc.: Natick, Massachusetts.

29. RStudio Team (2021). RStudio: Integrated Development for R. RStudio, PBC, Boston, MA.

30. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using lme4. J Stat Softw 2015;67:1–48.

31. Smith JA, Ashurst HL, Jack S, Woodcock AA, Earis JE. The description of cough sounds by healthcare professionals. Cough 2006;2:1–9.

32. Song WJ, Hui CKM, Hull JH, et al. Confronting COVID-19-associated cough and the post-COVID syndrome: role of viral neurotropism, neuroinflammation, and neuroimmune responses. Lancet Respir Med 2021;9:533–544.

33. Pramono RXA, Imtiaz SA, Rodriguez-Villegas E. A cough-based algorithm for automatic diagnosis of pertussis. PLoS One 2016;11:1–20.

34. Kosasih K, Abeyratne UR, Swarnkar V, Triasih R. Wavelet Augmented Cough Analysis for Rapid Childhood Pneumonia Diagnosis. IEEE Trans Biomed Eng 2015;62:1185–1194.

35. Al-Khassaweneh M, Bani Abdelrahman R. A signal processing approach for the diagnosis of asthma from cough sounds. J Med Eng Technol 2013;37:165–171.

36. Magni C, Chellini E, Lavorini F, Fontana GA, Widdicombe J. Voluntary and reflex cough: Similarities and differences. Pulm Pharmacol Ther 2011;24:308–311.

37. Fontana GA, Widdicombe J. What is cough and what should be measured? Pharmacol Ther 2007;20:307.
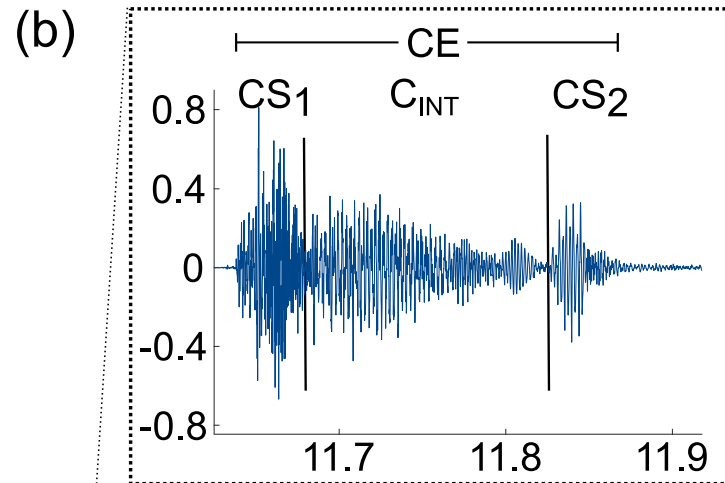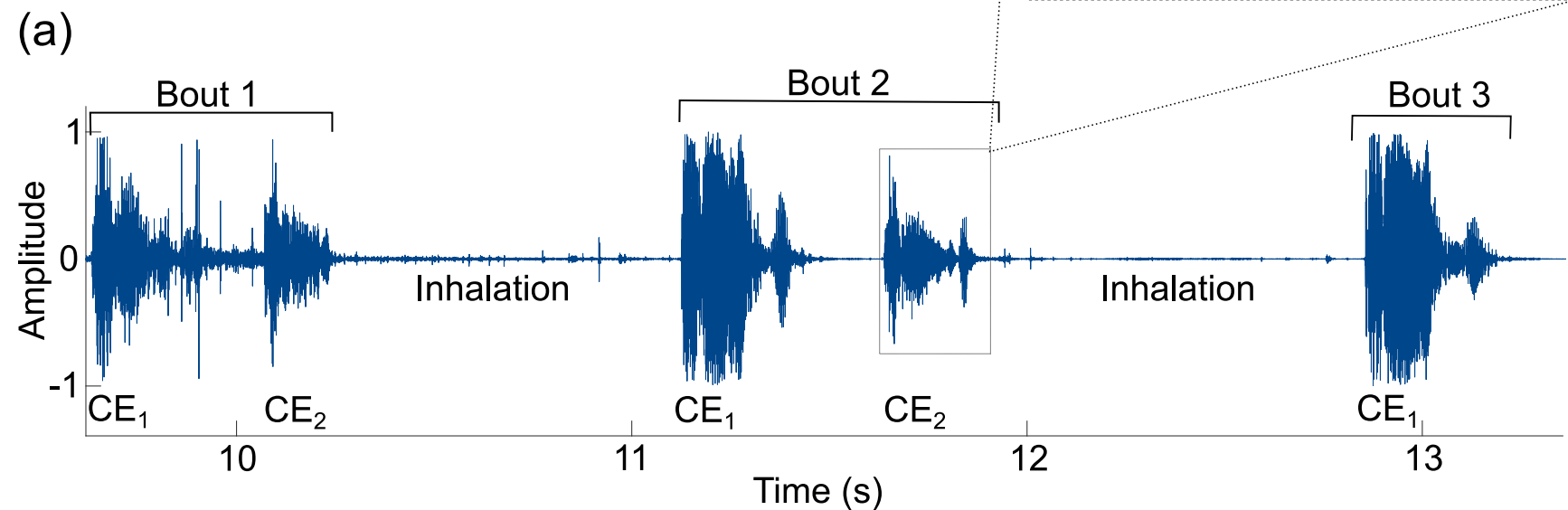
Figure 1: (a) Example raw signal recorded from a study participant with coronavirus disease 2019 (COVID-19) showing the classification of cough efforts (CEs), positions, and bouts. (b) The zoomed in portion of the figure shows the three constituent parts of a cough effort.
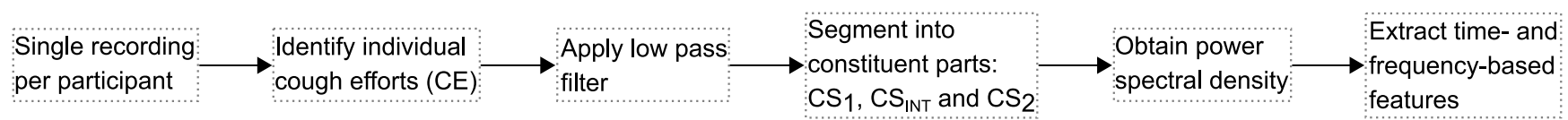
Figure 2: Flowchart showing the key analysis methods applied to the cough sound recordings from study participants.
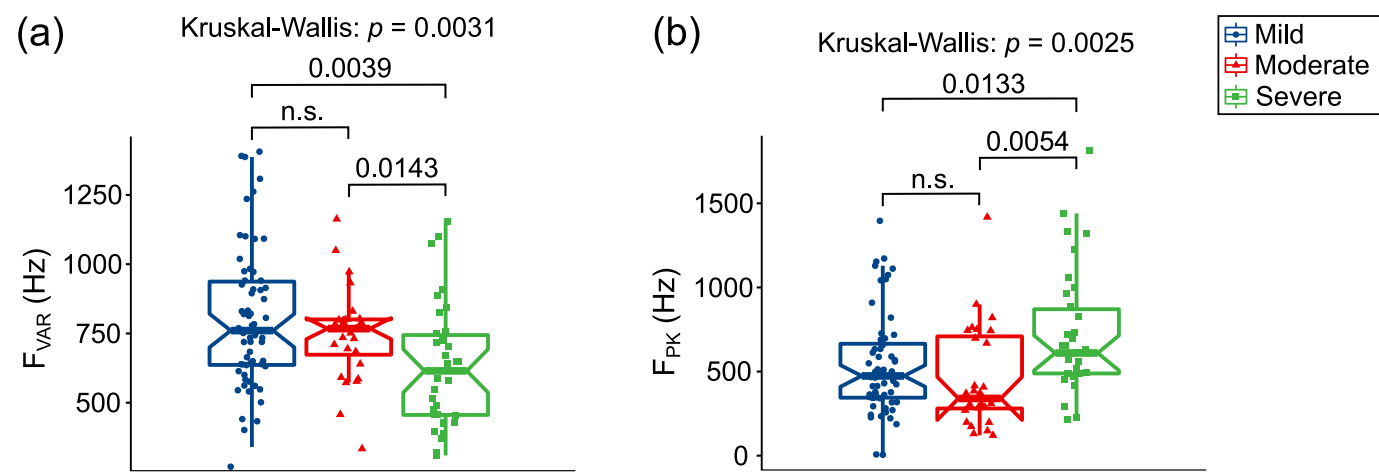
Figure 3: Boxplots showing (a) the frequency variability (F_{VAR}) and (b) the peak frequency (F_{PK}) of the whole cough effort (CE) signal for all patients and for different levels of disease severity. Each datapoint represents the mean value of that parameter in coughs at either position one (CE_1) or position 2 (CE_2).
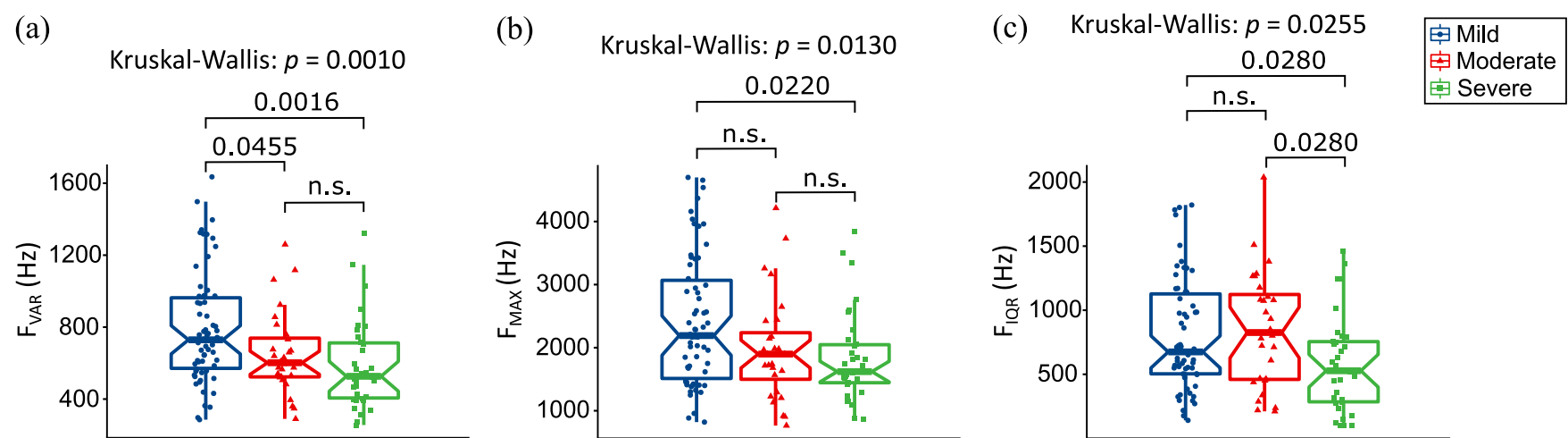
Figure 4: Boxplots showing (a) frequency variability (F_{VAR}) and (b) the frequency at maximum energy (F_{MAX}) of the first cough sound (CS_1) and (c) the interquartile range of frequency (F_{IQR}) of the intermediate part of the cough (C_{INT}) for all patients for different levels of disease severity. Each datapoint represents the mean value of that parameter in cough efforts at either position one (CE_1) or two (CE_2).
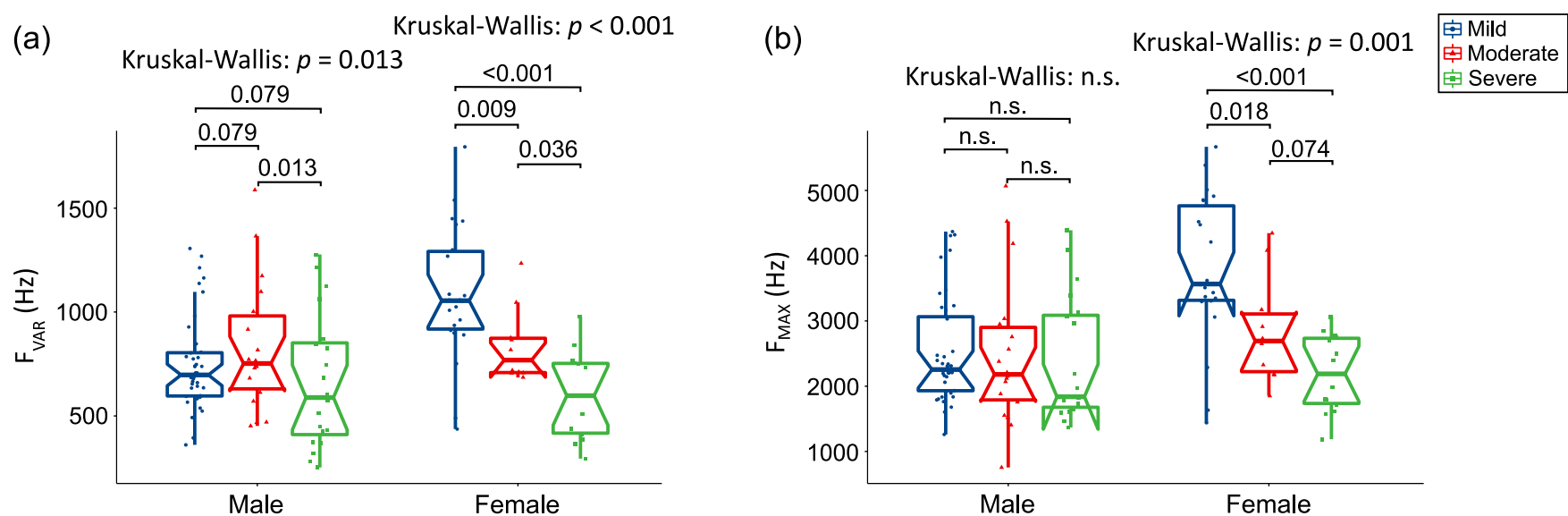
Figure 5: Boxplots showing (a) frequency variability (F_{VAR}) and (b) the frequency at maximum energy (F_{MAX}) of the intermediate part of the cough sound (C_{INT}). A significant interaction between gender and disease severity was found, so the Kruskal-Wallis test was applied to data for each gender separately.