



Original Research

Lung ultrasonography for long-term follow-up of COVID-19 survivors compared to chest CT scan

Guido Giovannetti^a, Lucrezia De Michele^b, Michele De Ceglie^c, Paola Pierucci^{d,e,*},
 Alessandra Mirabile^c, Marco Vita^f, Vincenzo Ostilio Palmieri^g,
 Giovanna Elisiana Carpagnano^{d,e}, Arnaldo Scardapane^c, Carlo D'Agostino^b

^a Department of Biomedical Sciences and Human Oncology, University of Bari Medical School "Aldo Moro", Piazza Giulio Cesare 11, 70124, Bari, Italy

^b Cardio Thoracic and Vascular Department, University Hospital Policlinico of Bari, Piazza Giulio Cesare 11, 70124, Bari, Italy

^c DIM, Interdisciplinary Department of Medicine, Section of Diagnostic Imaging, University of Bari Medical School "Aldo Moro", Piazza Giulio Cesare 11, 70124, Bari, Italy

^d Cardiothoracic Department, Respiratory and Critical Care Unit Bari Policlinic University Hospital, Italy

^e "Aldo Moro" Bari University School of Medicine, Italy

^f Ammagamma Artificial Intelligence Solutions, Via S. Orsola 37, 41121, Modena, Italy

^g Department of Biomedical Sciences and Human Oncology, Clinica Medica "A. Murri", University of Bari "Aldo Moro", Piazza Giulio Cesare 11, 70124, Bari, Italy



ARTICLE INFO

Keywords:

COVID-19 follow-up
 Lung ultrasonography
 Chest computed tomography
 Interstitial lung disease
 ILD
 Hypoxemic acute respiratory failure

ABSTRACT

Background: While lung ultrasonography (LUS) has utility for the evaluation of the acute phase of COVID-19 related lung disease, its role in long-term follow-up of this condition has not been well described. The objective of this study is to compare LUS and chest computed tomography (CT) results in COVID-19 survivors with the intent of defining the utility of LUS for long-term follow-up of COVID-19 respiratory disease.

Methods: Prospective observational study that enrolled consecutive survivors of COVID-19 with acute hypoxemic respiratory failure (HARF) admitted to the Respiratory Intensive Care Unit. Three months following hospital discharge, patients underwent LUS, chest CT, body plethysmography and laboratory testing, the comparison of which forms the basis of this report.

Results: 38 patients were enrolled, with a total of 190 lobes analysed: men 27/38 (71.1%), mean age 60.6 y (SD 10.4). LUS findings and pulmonary function tests outcomes were compared between patients with and without ILD, showing a statistically significant difference in terms of LUS score (p: 0.0002), FEV1 (p: 0.0039) and FVC (p: 0.012). ROC curve both in lobe by lobe and in patient's overall analysis revealed an outstanding ILD discrimination ability of LUS (AUC: 0.94 and 0.95 respectively) with a substantial Cohen's coefficient (K: 0.74 and 0.69).

Conclusions: LUS has an outstanding discrimination ability compared to CT in identifying an ILD of at least mild grade in the post COVID-19 follow-up. LUS should be considered as the first-line tool in follow-up programs, while chest CT could be performed based on LUS findings.

1. Introduction

Corona Virus Disease 2019 (COVID-19) pandemic is the major current global health concern, due to its high rate of hypoxemic acute respiratory failure (HARF) and the number of related deaths worldwide [1].

The long-term complications of COVID-19 pneumonia are starting to emerge but data from previous coronavirus outbreaks, such as SARS and MERS, suggest that some patients could experience long-term pulmonary complications such as ILD and pulmonary vascular disease [2].

Wang et al. [3] reported that 94% of the patients with COVID-19 pneumonia had residual CT findings after a median time from discharge of 25 days and ground glass opacities (GGO) have been identified as the most frequent residual pattern. On the contrary, the crazy-paving pattern was no longer observed after 14 days from the onset of initial symptoms, likely as a result of recovery [4].

Lung ultrasonography has demonstrated utility for management of COVID-19 ([5–9]) and gives similar results as chest CT for the evaluation of lung involvement ([10,11]). Its ease of use, low cost, and lack of radiation has led to its well-defined use during the COVID-19 pandemic. The role of LUS in the recovery phase of COVID-19 following hospital

* Corresponding author. Cardiothoracic Department, Respiratory and Critical care Unit Bari Policlinic University Hospital, Italy.

E-mail address: paola.pierucci@policlinico.ba.it (P. Pierucci).

Abbreviations

| | | | |
|----------|---|-------|-------------------------------------|
| AUC | (area under the curve) | LLDAM | (late lung damage) |
| COVID-19 | (CoronaVirus Disease-2019) | LLL | (left lower lobe) |
| CT | (computed tomography) | LUL | (left upper lobe) |
| CTEPH | (chronic thromboembolic pulmonary hypertension) | LUS | (lung ultrasonography) |
| CTPA | (computed tomography pulmonary angiography) | NPV | (negative predictive value) |
| FN | (false negative) | NRS | (non-invasive respiratory support) |
| FP | (false positive) | PE | (pleural effusion) |
| GGO | (ground glass opacity) | PFTs | (pulmonary function tests) |
| HARF | (hypoxemic acute respiratory failure) | PPV | (positive predictive value) |
| HRCT | (high resolution computed tomography) | RLL | (right lower lobe) |
| ILD | (interstitial lung disease) | RML | (right middle lobe) |
| K | (kappa) | ROC | (receiver operating characteristic) |
| | | RUL | (right upper lobe) |
| | | TSS | (total severity score) |

discharge has not yet been defined. To the best of our knowledge, this was the first long-term follow-up including a comparison between chest CT and LUS, performed at the same time after at least 90 days after discharge at home.

The aims of the present study were first, to assess the reliability of the LUS compared to the gold standard chest CT in detecting the presence of ILD in patients survived to COVID-19 with HARF, and second, to identify the LUS role in the long-term follow-up of these patients.

2. Methods

2.1. Study characteristics

This was a single centre prospective observational cohort study which followed the STROBE criteria for observational studies. From May 18th to July 25th, 2020 it was carried out among different Clinical Units of the Bari Policlinic University Hospital, being part of the post COVID-19 Late Lung Damage (LLDAM) follow-up study. It was approved by the Ethic Committee of University Hospital Policlinico of Bari (study number 6380, 12th May 2020) and all patients involved signed an informed and written consent before being enrolled.

2.2. Patients' characteristics

Patients included in this follow-up study had the following characteristics at the time of hospitalization:

- patients admitted to Respiratory Intensive Care Unit (RICU);
- age between 18 and 75 years;
- positive RT-PCR testing for COVID-19 on nasopharyngeal swabs;
- severe disease's onset with at least one of the following criteria: HARF with respiratory rate >30 breaths/min, SpO₂ <93% on room air in resting position, PaO₂/FiO₂ <300 mmHg, requiring FiO₂>60% and non-invasive respiratory support (NRS); septic shock; multiple organ failure;
- no history of previous pulmonary fibrosis or pulmonary hypertension disease.

2.3. Clinical and instrumental evaluation

Two negative swabs and serology testing were required prior to be evaluated for the present study.

LLDAM protocol provided cardiologic and pulmonary clinical and instrumental evaluation. Chest CT, LUS and laboratory tests were performed within 3 months from admission in hospital in order to evaluate the late pulmonary and cardiovascular infection-related sequelae. In more detail, a Computed Tomography Pulmonary Angiography (CTPA) was performed to detect possible signs of chronic thromboembolic

pulmonary hypertension (CTEPH), as per the current literature [12]. Simultaneously, High Resolution Computed Tomography (HRCT) was performed to identify potential presence of ILD.

In case of persistent clinical symptoms, instrumental abnormalities (ILD) and pulmonary function impairment, an appropriate respiratory evaluation and treatment plan was initiated to allow a complete lung recovery.

The CT scans were obtained with a 128 row multi-detector CT (Siemens Somatom Definition DS). An unenhanced scan in supine position from the jugular to the diaphragmatic domes was performed, followed by a CTPA. Images were reconstructed with a slice-thickness of 1 mm in mediastinal and parenchymal windows. The CT images were independently reviewed by two radiologists (M.D.C. and A.M., senior specialist consultant and fellow respectively) and final decision was reached by consensual discussion. LUS examinations were carried out via Philips IE 33 Ultrasound System with linear array probe (MHz 7.5–10) Image depth was adjusted at 4–6 cm and a single-focal point modality was used with the focal point set on the pleural line. Tissue Harmonics Imaging was disabled. Gain was adjusted throughout the examination to optimize the image. Every thoracic region was evaluated both in transverse and longitudinal scans, with patient seated. The LUS images were independently reviewed by two operators (L.D.M. and G. G., senior specialist consultant and fellow respectively) and final decision was reached by consensual discussion.

2.4. LUS and chest CT grading

LUS evaluation recorded the presence or absence of the following findings: A-lines, B-lines, a thickened pleural line, consolidations and pleural effusion. These features have been well described in COVID-19 patients in previous studies [13]. Aeration pattern was evaluated on LUS as per the current literature [14]: score 0 as normal pattern, A-lines or <3 B-lines (Fig. 1A); 1 as moderate loss, ≥3 B-lines (Fig. 1B); 2 as severe loss, coalescent B-lines (Fig. 1C); 3 as complete loss, white lung (Fig. 1D) and/or lung consolidations. The total LUS score was the sum of the points from each lobe and ranges from 0 to 36 points.

All these findings were evaluated and graded and then compared to CT features. For the purpose of the study, HRCT images were analysed to identify eventual presence of GGO and intra-lobular interstitial thickening. The quantitative involvement of each lobe was recorded awarding a CT score from 0 to 5, depending on a visual assessment of the percentage of the parenchymal involved, as per the current literature [15]: score 0 as no involvement; 1 as < 5%; 2 as 5%–25%; 3 as 26%–49%; 4 as 50%–75%; 5 as > 75%. The Total Severity Score (TSS) was the sum of the points from each lobe and ranges from 0 to 25 points. Additionally, eventual presence of consolidations and PE was recorded.

In order to line up the ultrasonography vision the closest possible to the chest CT scan, the topography of the thorax was divided into 12

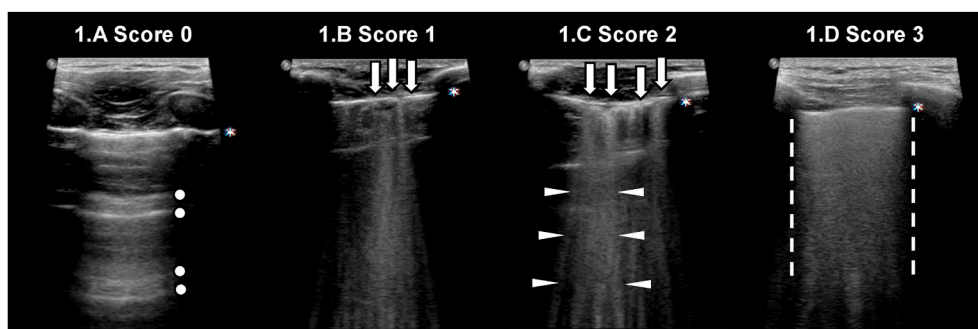


Fig. 1. Lung ultrasonography grading score. Score 0: normal pattern, A-lines or <3 B-lines (Fig. 1A); score 1: moderate loss, ≥ 3 B-lines (Fig. 1B); score 2: severe loss, coalescent B-lines (Fig. 1C); score 3: complete loss, white lung (Fig. 1D) and/or lung consolidations. Legend: pleural line is indicated by asterisk; A-lines are indicated by circles; B-lines are indicated by arrows; coalescent B-lines are indicated by triangles; white lung is indicated by dashed lines.

regions and grouped according to the lung's lobes, according to Yang et al. [6]:

- Right upper lobe (RUL) corresponding to R1 and R3;
- Right middle lobe (RML) corresponding to R2;
- Right lower lobe (RLL) corresponding to R4, R5 and R6;
- Left upper lobe (LUL) corresponding to L1, L2 and L3;
- Left lower lobe (LLL) corresponding to L4, L5 and L6.

2.5. Statistical analysis

The diagnostic performance of LUS was evaluated through receiver operating characteristic (ROC) curve analysis [16], quantifying the accuracy of LUS in discriminating between patients with and without ILD, as assessed via chest CT. The area under the ROC curve (AUC) of 0.5 is considered as no discrimination ability, AUC 0.5–0.7 poor, 0.7–0.8 acceptable, 0.8–0.9 excellent, and ≥ 0.9 outstanding [17].

Cohen's kappa (K) test was also run to establish the LUS ability to assess severity of ILD, comparing to CT findings. The Cohen's K is a coefficient representing the agreement between two methods. It is rated as follows: values ≤ 0 as indicating no agreement; 0.01–0.20 slight; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; 0.81–1.00 almost perfect agreement [18].

Ultrasonographic and radiographic findings were analysed both considering the final score of each method for each patient (LUS score and TSS) and lobe by lobe, thus to increase the number of cases compared. Presence of disease was defined for CT and LUS score cut-off of 1.

Further statistical analysis was performed in order to confirm LUS ability to identify ILD affecting <5% in each lobe. This was reached applying a mixed cut-off which considered a CT score of 1 (i.e. <5% lobe parenchymal involvement) both in the case of minimal presence or in absence of disease at LUS.

Sunburst diagram was also used for visual representation of CT and LUS concordance. It shows the hierarchical relation between the central circle which represents the CT scan findings and the outer ring which represents the LUS findings.

Chest CT and LUS outcomes were compared using χ^2 test. Pulmonary function tests (PFTs) outcomes were compared in patients with and without signs of ILD using student's *t*-test and χ^2 test between variables. *P*-value <0.05 was considered statistically significant.

3. Results

Characteristics and demographic of patients enrolled are shown in Table 1. The study enrolled 38 patients, a total of 190 lung lobes analysed.

Main imaging and pulmonary function tests outcomes are shown in Table 2. On LUS examination, 14/38 (36.8%) patients did not show any

Table 1

Patients' demographic and characteristics. of patients included, N = 38.

| SEX, male, N (%) | 27 (71.1%) | |
|---------------------------------------|-----------------|---------------|
| AGE, y, mean (SD) | 60.6 (10.4) | |
| BMI, kg/m ² , mean (SD) | 27.5 (4.1) | |
| COMORBIDITY, N (%) | 34 (89.5%) | |
| HYPERTENSION, N (%) | 21 (55.3%) | |
| DYSLIPIDAEMIA, N (%) | 8 (21.1%) | |
| OBESITY, N (%) | 7 (18.4%) | |
| DIABETES MELLITUS, N (%) | 6 (15.8%) | |
| CKD, N (%) | 6 (15.8%) | |
| CHRONIC CHD, N (%) | 5 (13.2%) | |
| MALIGNANCY, N (%) | 5 (13.2%) | |
| THYROID DISEASE, N (%) | 4 (10.5%) | |
| OSAS, N (%) | 3 (7.9%) | |
| ASTHMA, N (%) | 1 (2.6%) | |
| COPD, N (%) | 1 (2.6%) | |
| SMOKER, N (%) | 18 (47.4%) | |
| CURRENT SMOKER, N (%) | 2 (5.3%) | |
| EX-SMOKER, N (%) | 16 (42.1%) | |
| CLINICAL SITUATION AT HOSPITALIZATION | | |
| HARF, N (%) | 38 (100%) | |
| ARDS, N (%) | 25 (65.8%) | |
| MILD (P/F: 200–300), N (%) | 11 (28.9%) | |
| MODERATE (P/F: 100–199), N (%) | 11 (28.9%) | |
| SEVERE (P/F: <100), N (%) | 3 (7.9%) | |
| NRS, N (%) | 38 (100%) | |
| CPAP, N (SD) | 26 (68.4%) | |
| HFNC, N (SD) | 8 (21.1%) | |
| BIPAP, N (SD) | 4 (10.5%) | |
| LABORATORY TESTS COMPARISON | | |
| CPR, mg/dl (SD) | 106.4 (66.9) | 3.3 (1.4) |
| D-DIMERS, ug/ml (SD) | 1345.5 (1542.5) | 303.9 (220.9) |
| NEUTROPHILS, % (SD) | 77.1 (8.9) | 55.4 (7.9) |
| LYMPHOCYTES, % (SD) | 15.6 (7.0) | 35.0 (7.2) |

Abbreviations BMI: Body Mass Index; CKD: Chronic Kidney Disease; CHD: Coronary Heart Disease; OSAS: Obstructive Sleep Apnea Syndrome; COPD: Chronic obstructive pulmonary disease; HARF: hypoxemic acute respiratory failure; ARDS acute respiratory distress syndrome; NRS: non-invasive respiratory support; CPAP: Continuous Positive Airway Pressure; HFNC: High Flow Nasal Cannula; BiPAP: Bilevel Positive Airway Pressure; CPR: C-reactive Protein.

sign of ILD while 24/38 (63.2%) did, visualised via B-lines. On CT examination, 16/38 (42.1%) patients showed no signs of ILD while 22/38 (57.9%) patients showed signs ILD. Among the latter, 14/38 (36.8%) showed GGO, 6/38 (15.8%) showed GGO plus sub-pleural bands and 2/38 (5.3%) showed intra-lobular interstitial thickening. All patients with signs of ILD (22/38, 57.9%) and pulmonary function test impairment (reduced FEV1 and FVC) underwent a full respiratory consultation and, if required, a further treatment plan.

None of the patients included in our study population had pleural effusion nor were consolidations on LUS examination, and these data were confirmed on chest CT scan.

According to Yang et al. [6], ultrasound scanning areas were grouped

Table 2
Imaging and PFTs outcomes of patients included, N = 38.

| IMAGING OUTCOMES | | | |
|-----------------------------------|-----------------|------------------|---------|
| | CT | LUS | P value |
| NO SIGNS OF ILD, N (%) | 16 (42.1%) | 14 (36.8%) | 0.75 |
| SIGNS OF ILD, N (%) | 22 (57.9%) | 24 (63.2%) | 0.81 |
| MILD, N (%) | 15 (39.5%) | 16 (42.1%) | 0.87 |
| MODERATE, N (%) | 7 (18.4%) | 8 (21.1%) | 0.81 |
| SEVERE, N (%) | 0 (0.0%) | 0 (0.0%) | - |
| CONSOLIDATIONS, N (%) | 0 (0.0%) | 0 (0.0%) | - |
| PLEURAL EFFUSION, N (%) | 0 (0.0%) | 0 (0.0%) | - |
| PULMONARY FUNCTION TESTS OUTCOMES | | | |
| | No ILD* | ILD [†] | P value |
| PH (SD) | 7.43 (0.03) | 7.43 (0.04) | 0.88 |
| PACO2, mmHg (SD) | 39.88 (3.30) | 41.32 (8.58) | 0.53 |
| PAO2, mmHg (SD) | 86.13 (9.76) | 83.68 (13.40) | 0.54 |
| SPO2, % (SD) | 96.56 (1.15) | 96.18 (1.37) | 0.37 |
| P/F RATIO, mmHg (SD) | 411.19 (45.84) | 401.68 (54.36) | 0.57 |
| 6MWT, mt, (SD) | 528.25 (123.09) | 509.68 (65.77) | 0.55 |
| BORG PRE-TEST (SD) | 0.50 (0.52) | 0.68 (0.95) | 0.49 |
| BORG POST-TEST (SD) | 3.25 (1.48) | 3.23 (1.45) | 0.96 |
| HR PRE-TEST, bpm (SD) | 75.75 (9.49) | 72.27 (10.88) | 0.31 |
| HR POST-TEST, bpm (SD) | 101.56 (13.71) | 93.32 (11.56) | 0.05 |
| SPO2 PRE-TEST, % (SD) | 97.13 (1.45) | 97.41 (1.00) | 0.48 |
| SPO2 POST-TEST, % (SD) | 96.38 (1.20) | 96.50 (1.06) | 0.74 |
| FVC, L (SD) | 4.15 (0.91) | 3.35 (0.86) | 0.01 |
| FVC, % of predicted (SD) | 100.86 (16.01) | 97.82 (16.27) | 0.59 |
| FEV1, L (SD) | 3.40 (0.65) | 2.74 (0.61) | <0.01 |
| FEV1, % of predicted (SD) | 102.14 (14.82) | 100.82 (14.15) | 0.79 |
| FEV1/FVC, % (SD) | 78.80 (3.20) | 79.70 (5.05) | 0.56 |
| TLC, L (SD) | 6.35 (1.44) | 5.62 (1.16) | 0.10 |
| TLC, % of predicted (SD) | 97.15 (17.74) | 93.34 (12.77) | 0.46 |
| DLCO, mL/min/mmHg (SD) | 86.27 (13.85) | 87.53 (12.20) | 0.79 |
| DLCO/VA, % of predicted (SD) | 95.82 (11.12) | 102.51 (13.58) | 0.16 |

P values statistically significant are in bold font (p value ≤ 0.05). No signs of ILD corresponding to CT score = 0 and LUS score = 0. Signs of ILD corresponding to CT score = 1–25 and LUS score = 1–36. Mild ILD corresponding to CT score = 1–5 and LUS score = 1–7. Moderate ILD corresponding to CT score = 6–12 and LUS score = 8–18. Severe ILD corresponding to CT score 13–25 and LUS score 19–36. *No ILD is defined as CT score = 0. [†]ILD is defined as CT score ≥ 1 . Abbreviations: PFTs: Pulmonary function tests; ILD: interstitial lung disease; TSS: total severity score; LUS: lung ultrasound; PaCO2: partial arterial pressure of carbon dioxide; PaO2: partial arterial pressure of oxygen; SpO2: peripheral capillary oxygen saturation; P/F ratio: partial arterial pressure of oxygen/fraction of inspired oxygen ratio; 6MWT: 6 min walking test; HR: heart rate; FVC: forced vital capacity; FEV1: forced expiratory volume in 1 s; TLC: total lung capacity; DLCO: Diffusing capacity of the lung for carbon monoxide; VA: alveolar volume.

to identify corresponding lobes, thus to perform a lobe by lobe analysis between LUS and CT findings, with cut-off value of 1 (Fig. 2A). ROC curves revealed an AUC of 0.77 (acceptable correlation) and the Cohen's K coefficient was 0.46 (moderate agreement rate). Among false negatives (FN, 29/105 27.6%), 27/29 had minimal ILD on CT (CT score = 1, i.e. 1–5% lung involvement) and 2/29 had at least a mild ILD (CT score ≥ 2 , i.e. >5% lung involvement) located in the RML in both two cases. Fig. 2C shows LUS findings considering the overall evaluation of the 38 patients with a CT cut-off value of 1. ROC curves revealed an AUC of 0.82 (excellent discrimination ability) and the Cohen's K coefficient was 0.45 (moderate correlation). Among the FN cases (4/14, 28.6%), none of the patients had a TSS score >5 (i.e. at least a mild ILD in one or more lobes).

Conversely, applying the mixed cut-off (Fig. 2B), the AUC in the lobe by lobe analysis raised to 0.94 (outstanding discrimination ability) and Cohen's K coefficient raised to 0.74 (substantial correlation). Similarly, in the comparison between the TSS and LUS score in the patient's overall analysis (Fig. 2D) the AUC raised to 0.95 (outstanding discrimination ability) and Cohen's K coefficient raised to 0.69 (substantial correlation).

A good visual representation of CT and LUS concordance was

reached through the Sunburst chart (e-figure). In the patient's overall analysis, the presence of a TSS score >5 correlated with high LUS score in each patient. In the lobe by lobe analysis, Sunburst diagrams show that CT score ≥ 2 matched the presence of disease in the LUS scoring in each single lobe as represented in LUL, with the only exception of RML (where in 2/105 (1.9%) cases no disease was found on LUS whereas it was present at the CT scan (e-figure).

Table 3 shows the comparison between mean LUS total score and each lobe score in patients with and without ILD, as assessed via chest CT. P values were strongly statistically significant ($p < 0.05$) in LUS score and in each lobe comparison, with the only exception of RML.

Fig. 3B1 shows that in the both lungs analysis ROC curve revealed an AUC of 0.83 and Cohen's K coefficient was 0.54 with an excellent discrimination ability and a moderate agreement rate. Considering the left lung individually (Fig. 3B2), AUC was 0.86 and Cohen's K was 0.63 with an excellent discrimination ability and a substantial agreement rate. On the other hand, in the right lung (Fig. 3B3), AUC was 0.80 and Cohen's K was 0.45 with an excellent discrimination ability and a moderate agreement rate.

4. Discussion

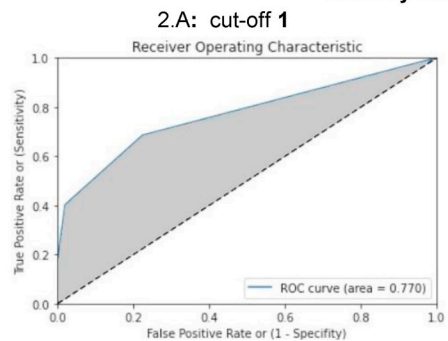
This study results showed that in COVID-19 survivors with severe lung involvement and HARF, the ILD findings identified through the LUS were confirmed by the CT scan with an outstanding discrimination ability and a substantial agreement rate. Hence, the study results indicate that LUS may serve as an effective first line tool for the longitudinal assessment of COVID-19 patients cared for Intensive Care Unit following hospital discharge, as proposed by George et al. [2] in a recent state of art review.

In consideration of both the lobe by lobe and the final score of the patient analysis, the possible explanation for the FP percentage may be found in the time interval between LUS and CT scan performance. Indeed, among the 6/38 patients with a LUS score ≥ 1 , who showed no residual ILD at CT scan, 3/38 patients had LUS performed at least 26 days before the CT scan for technical and logistical reasons pandemic related, so patients could probably have further recovered from ILD over this period of time. For the remaining 3/38 cases, the evaluation could probably have been affected by artefacts (Z-lines, comet tail or other common artefacts [19]) or bad acoustic window of the patient. Optimal ultrasound propagation may be prevented in obese or overweight patients (18.4% and 50% of our study population respectively), due to thickness of their ribcage and soft tissues, or on the other hand, in patients with reduced intercostal space (16% of our study population).

The FN may be related to LUS inability to identify minimal ILD (CT score = 1). Indeed, in favourable conditions (i.e. absence of artefacts and good acoustic window), LUS can be able to assess the presence of a minimal ILD, otherwise LUS cannot exclude the presence of a minimal ILD. In order to prove this hypothesis, a mixed CT cut-off to the Cohen concordance index was applied considering a CT score of 1 both in case of presence of minimal or in absence of disease at LUS. By using this mixed cut-off, the Cohen's K coefficient in lobe by lobe analysis raised from 0.46 to 0.74 (substantial correlation) with an outstanding discrimination ability (AUC: 0.94) and an improvement of all remaining parameters. Remarkably, FN dropped from 27.6% (29/105) to 1.9% (2/105). The lobe involved in both FN cases was the RML, which is the most difficult lobe to identify on LUS. Applying a mixed cut-off to the final score of each method (LUS score and TSS), Cohen's K coefficient raised from 0.45 to 0.69 (substantial agreement rate) with an outstanding discrimination ability (AUC: 0.95) and an improvement of all remaining parameters. With the mixed cut-off score the FN dropped from 28.6% to 0%, confirming the value of LUS in selecting FN only among those patients with real lower ILD involvement $<5\%$ /lobe.

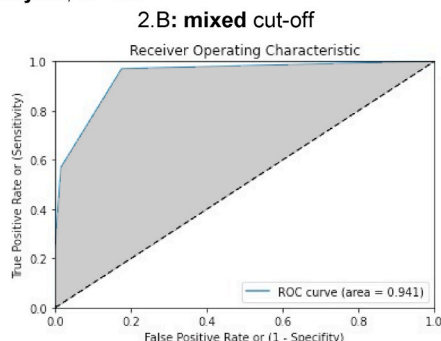
This study has several limitations. First, the small number of participants enrolled, however, in order to enhance the precision and reliability of the comparison between the two methods, a lobe by lobe

Lobe by lobe analysis, N=190



| | ILD | No ILD | |
|---------|--------------|--------------|-------------------------|
| LUS ≥ 1 | 63 | 22 | PPV: 74.1% FP: 25.9% |
| LUS = 0 | 76 | 29 | NPV: 72.4% FN: 27.6% |
| | Sens.: 68.5% | Spec.: 77.6% | |

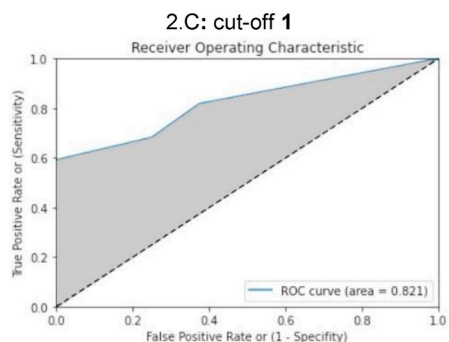
K: 0.46 Acc.: 73.2% LR⁺: 3.06 LR⁻: 0.41



| | ILD | No ILD | |
|---------|--------------|--------------|-------------------------|
| LUS ≥ 1 | 63 | 22 | PPV: 74.1% FP: 25.9% |
| LUS = 0 | 103 | 2 | NPV: 98.1% FN: 1.9% |
| | Sens.: 96.9% | Spec.: 82.4% | |

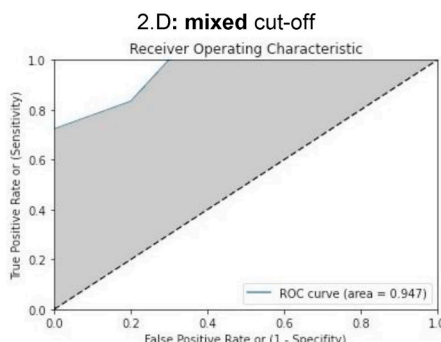
K: 0.74 Acc.: 87.4% LR⁺: 5.51 LR⁻: 0.04

Patient's overall analysis, N=38



| | ILD | No ILD | |
|---------|--------------|--------------|-------------------------|
| LUS ≥ 1 | 18 | 6 | PPV: 75.0% FP: 25.0% |
| LUS = 0 | 10 | 4 | NPV: 71.4% FN: 28.6% |
| | Sens.: 81.8% | Spec.: 62.5% | |

K: 0.45 Acc.: 73.7% LR⁺: 2.18 LR⁻: 0.29



| | ILD | No ILD | |
|---------|---------------|--------------|-------------------------|
| LUS ≥ 1 | 18 | 6 | PPV: 75.0% FP: 25.0% |
| LUS = 0 | 14 | 0 | NPV: 100.0% FN: 0.0% |
| | Sens.: 100.0% | Spec.: 70.0% | |

K: 0.69 Acc.: 84.2% LR⁺: 3.33 LR⁻: 0.00

Table 3

LUS findings in patients with and without ILD, as assessed via chest CT.

| | No ILD* | ILD [†] | P value |
|----------------------------|-------------|------------------|----------------|
| LUS TOTAL SCORE, mean (SD) | 1.13 (1.59) | 5.50 (3.97) | 0.00019 |
| RUL, mean (SD) | 0.32 (0.58) | 1.21 (0.78) | 0.00032 |
| RLL, mean (SD) | 0.28 (0.46) | 1.50 (1.28) | 0.00048 |
| RML, mean (SD) | 0.26 (0.45) | 0.40 (0.51) | 0.38077 |
| LUL, mean (SD) | 0.17 (0.51) | 1.50 (1.05) | 0.00002 |
| LLL, mean (SD) | 0.20 (0.41) | 1.56 (1.38) | 0.00017 |

P values statistically significant are in bold font (p value ≤ 0.05). *No ILD corresponding to CT score = 0; [†]ILD corresponding to CT score ≥ 1. Abbreviations: LUS: Lung Ultrasonography Score; ILD: Interstitial Lung Disease; CT: Computed Tomography; RUL: Right Upper Lobe; RLL: Right Lower Lobe; RML: Right Middle Lobe; LUL: Left Upper Lobe; LLL: Left Lower Lobe.

analysis was also performed, thus increasing the number of cases compared from 38 patients to 190 single lobes. Second, LUS findings can be subject to inter-observer bias due to operator's skill however, it is easy-to-learn, less technically demanding than other sonographic examinations [20] and all exams were reviewed by an experienced senior specialist consultant. Third, LUS cannot correctly identify RML because

Fig. 2. Lobe by lobe and patient's overall analysis. ROC and Cohen's kappa analysis carried out with a CT cut-off established of 1 (minimum interstitial disease included) in the lobe by lobe analysis (2.A) and in the patient's overall analysis (2.C). Same analysis carried out with a mixed cut-off (CT score of 1 included both in the case of minimal presence or in absence of disease at LUS) in the lobe by lobe analysis (2.B) and in the patient's overall analysis (2.D). Abbreviations: ROC: receiver operating characteristic; K: Cohen's kappa; Sens.: sensitivity; Spec.: specificity; PPV: positive predictive value; NPV: negative predictive value; Acc.: accuracy; LR⁺: Likelihood ratio positive; LR⁻: Likelihood ratio negative; FP: false positive; FN: false negative.

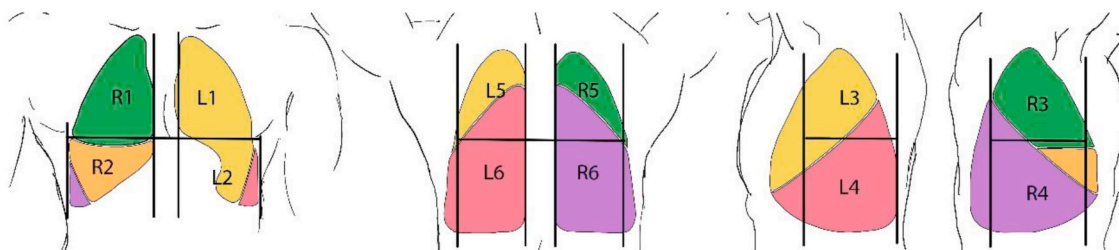
it is the smallest lobe and it may be subject to inter-individual variability due to possible ribcage conformation, diaphragmatic excursion and anatomical variations. Last, LUS is unable to identify minimal ILD, however, the patients should always undergo a full clinical evaluation to ensure the resolution of the clinical symptoms in correlation with the radiological findings.

On the other hand, the major strengths of this study are: first, it demonstrated that LUS may effectively substitute chest CT scan becoming the first line examination tool to define the presence of ILD in the follow-up of COVID-19 survivors. Second, by using LUS, serial examinations can be performed, without any long-term risks related to ionizing radiation, both in the outpatient ambulatory practise and at home. Lastly, the low cost, easy and quick availability, confirm LUS to be a first-line ideal tool in the long-term follow-up of COVID-19 patients, guiding the physician towards an effective treatment plan.

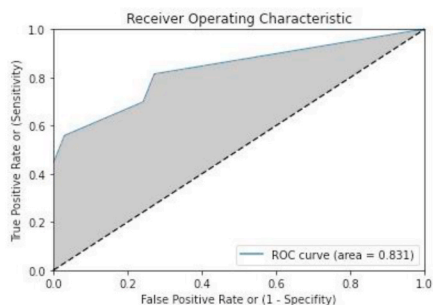
5. Conclusions

This study proved that LUS has an outstanding discrimination ability and a substantial agreement rate compared to the chest CT scan in the assessment and grading of ILD in patients at 3 months after COVID-19

3.A Thorax division into 12 regions and corresponding lobes



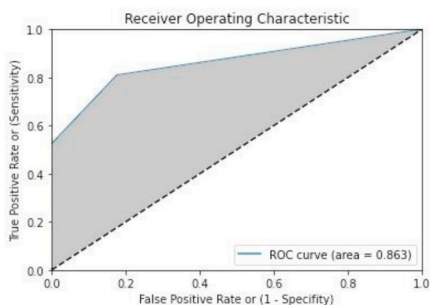
3.B.1: Both lungs analysis, N = 76



| | ILD | No ILD | |
|---------|-----------------|-----------------|-------------------------|
| LUS ≥ 1 | 35 | 9 | PPV: 79.5% FP: 20.5% |
| LUS = 0 | 24 | 8 | NPV: 75.0% FN: 25.0% |
| | Sens.: 81.4% | Spec.: 72.7% | |

K: 0.54 Acc.: 77.6% LR⁺: 2.98 LR⁻: 0.26

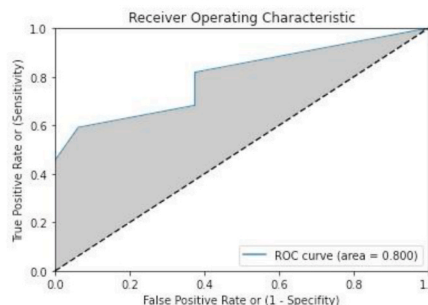
3.B.2: Left lung analysis, N = 38



| | ILD | No ILD | |
|---------|-----------------|-----------------|-------------------------|
| LUS ≥ 1 | 17 | 3 | PPV: 85.0% FP: 15.0% |
| LUS = 0 | 14 | 4 | NPV: 77.8% FN: 22.2% |
| | Sens.: 81.0% | Spec.: 82.4% | |

K: 0.63 Acc.: 81.6% LR⁺: 4.60 LR⁻: 0.23

3.B.3: Right lung analysis, N = 38



| | ILD | No ILD | |
|---------|-----------------|-----------------|-------------------------|
| LUS ≥ 1 | 18 | 6 | PPV: 75.0% FP: 25.0% |
| LUS = 0 | 10 | 4 | NPV: 71.4% FN: 28.6% |
| | Sens.: 81.8% | Spec.: 62.5% | |

K: 0.45 Acc.: 73.7% LR⁺: 2.18 LR⁻: 0.29

Fig. 3. Thorax division into 12 regions and corresponding lobes, with two lungs' analysis. 3.A: Thorax division into 12 regions. Each region was grouped according to the lung's lobes: right upper lobe (green) corresponding to R1 and R3; right middle lobe (orange) corresponding to R2; right lower lobe (violet) corresponding to R4, R5 and R6; left upper lobe (yellow) corresponding to L1, L2 and L3; left lower lobe (amaranth) corresponding to L4, L5 and L6. Analysis carried out with a CT cut-off established of 1 (minimum interstitial disease included) in the two lungs altogether (3.B.1), in the left lung (3.B.2) and in the right lung (3.B.3). Abbreviations: ROC: receiver operating characteristic; grey: area under the curve; dashed line: reference line; K: Cohen's kappa; Sens.: sensitivity; Spec.: specificity; PPV: positive predictive value; NPV: negative predictive value; Acc.: accuracy; LR⁺: Likelihood ratio positive; LR⁻: Likelihood ratio negative; FP: false positive; FN: false negative. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

severe lung infection with HARF. Therefore, LUS should be considered as the first-line tool in the follow-up of COVID 19 survivors looking for ILD, and it may guide the physician towards an effective treatment plan.

Further studies are required to provide evidence of the correct timing for LUS to be performed after discharge and how it may precisely influence the pharmacological therapy management.

Statement of ethics

This study was conducted ethically in accordance with the World Medical Association Declaration of Helsinki. It was approved by the

Ethic Committee of University Hospital Policlinico of Bari (study number 6380, 12th May 2020) and all patients involved signed an informed and written consent before being enrolled.

Funding sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

Dr De Michele and Dr Giovannetti had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: De Michele, Giovannetti.

Acquisition, analysis, or interpretation of data: De Michele, Giovannetti, De Ceglie, Mirabile, Pierucci.

Drafting of the manuscript: De Michele, Giovannetti, Pierucci.

Critical revision of the manuscript for important intellectual content:

All authors.

Statistical analysis: Vita.

Administrative, technical, or material support: All authors.

Supervision: V.O. Palmieri; G.E. Carpagnano; A. Scardapane; C. D'Agostino.

CRedit authorship contribution statement

Guido Giovannetti: Conceptualization, Methodology, Data curation, Writing – original draft. **Lucrezia De Michele:** Conceptualization, Methodology, Data curation, Writing – original draft. **Michele De Ceglie:** Data curation, Supervision. **Paola Pierucci:** Writing – original draft. **Alessandra Mirabile:** Data curation. **Marco Vita:** Software, Formal analysis. **Vincenzo Ostilio Palmieri:** Writing – review & editing. **Giovanna Elisiana Carpagnano:** Writing – review & editing. **Arnaldo Scardapane:** Writing – review & editing. **Carlo D'Agostino:** Writing – review & editing.

Declaration of competing interest

All the authors have no conflicts of interest to declare.

Acknowledgments

The authors would like to thank all the patients who participated in the study, Dr S. Cascella, Dr P. Colonna, Mrs Francesca Cagnetta, Mrs Anna Maria Caldarola and all physicians and nurses of the Cardiothoracic Department, Cardiovascular and Respiratory and Critical care unit of the Bari Policlinic University hospital.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rmed.2021.106384>.

References

- [1] COVID-19 Dashboard by the centre for systems science and engineering (CSSE) at Johns Hopkins University, Available at: <https://coronavirus.jhu.edu/map.html>, 2020. (Accessed 10 November 2020).

- [2] P.M. George, S.L. Barratt, R. Condliffe, R. Desai, A. Devaraj, I. Forrest, et al., Respiratory follow-up of patients with COVID-19 pneumonia, *Thorax* 75 (2020) 1009–1016.
- [3] Y. Wang, C. Dong, Y. Hu, C. Li, Q. Ren, X. Zhang, et al., Temporal changes of CT findings in 90 patients with COVID-19 pneumonia: a longitudinal study, *Radiology* 296 (2) (2020) E55–E64, <https://doi.org/10.1148/radiol.2020200843>.
- [4] F. Pan, T. Ye, P. Sun, S. Gui, B. Liang, L. Lingli, et al., Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19), *Radiology* 295 (3) (2020) 715–721, <https://doi.org/10.1148/radiol.2020200370>.
- [5] Y. Yang, Y. Huang, F. Gao, L. Yuan, Z. Wang, Lung ultrasonography versus chest CT in COVID-19 pneumonia: a two-centered retrospective comparison study from China, *Intensive Care Med.* 46 (9) (2020) 1761–1763, <https://doi.org/10.1007/s00134-020-06096-1>.
- [6] G. Lepri, M. Orlandi, C. Lazzeri, C. Bruni, M. Hughes, M. Bonizzoli, et al., The emerging role of lung ultrasound in COVID-19 pneumonia, *Eur. J. Rheumatol.* 7 (2020) S129–S133, <https://doi.org/10.5152/eurjrheum.2020.2063>.
- [7] M.J. Smith, S.A. Hayward, S.M. Innes, A. Miller, Point-of-care lung ultrasound in patients with COVID-19 - a narrative review, *Anaesthesia* 75 (2020) 1096–1104, <https://doi.org/10.1111/anae.15082>.
- [8] S. Sofia, A. Boccatonda, M. Montanari, M. Spampinato, D. D'Ardes, G. Cocco, et al., Thoracic ultrasound and SARS-COVID-19: a pictorial essay, *J. Ultrasound* 23 (2020) 217–221, <https://doi.org/10.1007/s40477-020-00458-7>.
- [9] A. Dargent, E. Chatelain, L. Kreitmann, J.P. Quenot, M. Cour, L. Argaud, Lung ultrasound score to monitor COVID-19 pneumonia progression in patients with ARDS, *PLoS One* 15 (7) (2020), e0236312, <https://doi.org/10.1371/journal.pone.0236312>.
- [10] A. Nouvenne, M.D. Zani, G. Milanese, A. Parise, M. Baciarello, E.G. Bignami, et al., Lung ultrasound in COVID-19 pneumonia: correlations with chest CT on hospital admission, *Respiration* 99 (2020) 617–624, <https://doi.org/10.1159/000509223>.
- [11] G. Soldati, A. Smargiassi, R. Inchingolo, D. Buonsenso, T. Perrone, D.F. Briganti, et al., Is there a role for lung ultrasound during the COVID-19 pandemic? *J. Ultrasound Med.* 39 (7) (2020) 1459–1462, <https://doi.org/10.1002/jum.15284>.
- [12] N. Galiè, M. Humbert, J.L. Vachiery, S. Gibbs, I. Lang, A. Torbicki, et al., ESC/ERS guidelines for the diagnosis and treatment of pulmonary hypertension: the joint task force for the diagnosis and treatment of pulmonary hypertension of the European society of cardiology (ESC) and the European respiratory society (ERS): endorsed by: association for European paediatric and congenital cardiology (AEPC), international society for heart and lung transplantation (ISHLT), *Eur. Heart J.* 37 (1) (2015) 67–119, <https://doi.org/10.1093/eurheartj/ehv317>, 2016.
- [13] Q.Y. Peng, X.T. Wang, L.N. Zhang, Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic, *Intensive Care Med.* 46 (2020) 849–850.
- [14] M.O. Ahmed, M. Reham, Diaphragmatic and lung ultrasound application as new predictive indices for the weaning process in ICU patients, *Egypt. J. Radiol. Nucl. Med.* 48 (1) (2017) 61–66, <https://doi.org/10.1016/j.ejrn.2017.01.005>.
- [15] P. Feng, Y. Tianhe, S. Peng, G. Shan, L. Bo, L. Lingli, et al., Time course of lung changes at chest CT during recovery from coronavirus disease 2019 (COVID-19), *Radiology* 295 (3) (2020) 715–721, <https://doi.org/10.1148/radiol.2020200370>.
- [16] J.A. Swets, Indices of discrimination or diagnostic accuracy: their ROCs and implied models, *Psychol. Bull.* 99 (1) (1986) 100–117.
- [17] A. Marques, S. Almeida, J. Carvalho, J. Cruz, A. Oliveira, C. Jacome, Reliability, validity, and ability to identify fall status of the balance evaluation systems test, mini-balance evaluation systems test, and brief-balance evaluation systems test in older people living in the community, *Arch. Phys. Med. Rehabil.* 97 (2016) 2166–2173, <https://doi.org/10.1016/j.apmr.2016.07.011>.
- [18] M.L. McHugh, Interrater reliability: the kappa statistic, *Biochem Med. Zagreb* 22 (3) (2012) 276–282.
- [19] M. Di Serafino, M. Notaro, G. Rea, F. Iacobellis, V. Delli Paoli, C. Acampora, et al., The lung ultrasound: facts or artifacts? In the era of COVID-19 outbreak, *Radiol. Med.* 125 (2020) 738–753, <https://doi.org/10.1007/s11547-020-01236-5>.
- [20] L. Gargani, G. Volpicelli, How I do it: lung ultrasound, *Cardiovasc. Ultrasound* 12 (2014) 25, <https://doi.org/10.1186/1476-7120-12-25>.