

Do Serial Lung Ultrasound Scores Predict Prolonged Mechanical Ventilation in Patients With Severe COVID-19? A Single-centre Retrospective Cohort Study

Hayato Taniguchi (✉ tanipan@yokohama-cu.ac.jp)

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center
<https://orcid.org/0000-0001-6995-3529>

Aimi Ohya

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center

Hidehiro Yamagata

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center

Masayuki Iwashita

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center

Takeru Abe

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center

Ichiro Takeuchi

Yokohama City University Medical Center: Yokohama Shiritsu Daigaku Fuzoku Shimin Sogo Iryo Center

Research

Keywords: coronavirus disease (COVID-19), lung ultrasound score, prolonged mechanical ventilation, inter-rater reliability, deep learning

Posted Date: June 16th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-614493/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Some patients with severe coronavirus disease (COVID-19) who present with fibrosis on computed tomography (CT) require prolonged mechanical ventilation (PMV). Lung ultrasound (LUS), a rapid, bedside test, has been reported to have findings consistent with those of CT. Thus, this study aimed to assess whether serial LUS scores could predict PMV or successful extubation in severe COVID-19 patients.

Methods: LUS was performed for 20 consecutive severe COVID-19 patients at three time points: admission (day 1), after 48 h (day 3), and seventh-day follow-up (day 7). We compared the LUS score with the results of chest X-rays and laboratory tests at three time points. Moreover, we assessed LUS score to determine the inter-rater reliability (IRR) of the results among examiners.

Results: While there were no significant differences in mortality in each PMV and successful extubation groups, there were significant differences in LUS scores on day 3 and day 7; XP score on day 7; and P/F ratio on day 7 in the PMV group ($p < 0.05$). There were significant differences in LUS scores on days 3 and 7, C-reactive protein (CRP) levels on day 7, and P/F ratio on day 7 in the successful extubation group ($p < 0.05$). The areas under the curves (AUCs) of LUS score on days 3 and 7, XP score on day 7, and P/F ratio were 0.88, 0.98, 0.77, and 0.80, respectively in the PMV group; and the AUCs of LUS score on days 3 and 7, CRP levels on day 7, and P/F ratio 0.79, 0.90, 0.82, and 0.79, respectively, in the successful extubation group. Variations in serial LUS scores exhibited significant differences between the groups. The serial LUS score on day 7 was higher than that on day 1 in the PMV group but lower in the successful extubation group ($p < 0.05$). However, there was slight IRR agreement in the LUS score changes on days 1 to 7 ($\kappa = 0.15$ [95% CI: 0-0.31]).

Conclusions: The serial LUS score of severe COVID-19 patients could predicted PMV and successful extubation. To overcome IRR disagreement, the automatic ultrasound judgement, such as deep learning, would be needed.

Background

Patients with prolonged mechanical ventilation (PMV) have a higher mortality rate and bear higher costs than those who do not require PMV [1]. The average duration of invasive mechanical ventilation in patients with coronavirus disease (COVID-19) admitted to the intensive care unit (ICU) was reported to be approximately 8.4 (95% confidence interval [CI] 1.6–13.7) days; however, in some patients, the use is prolonged [2, 3]. If patients receive PMV, they are usually excluded as candidates for extra-corporeal membranous oxygenation (ECMO), and with limited resources during a pandemic, it may be considered a withdrawal of treatment [4, 5]. The rapid surge of medical needs depletes ventilators and ICU beds, making the use of anaesthetic machines instead of ventilators compulsory [6]. Therefore, it is very important to predict whether patients will require PMV or can be extubated if ventilatory management becomes necessary.

It has been reported that patients with COVID-19 who require PMV present with fibrosis on computed tomography (CT) [7, 8]. Although CT is useful for assessing lung severity, it requires the transportation of critically ill, invasively ventilated patients to the radiology facilities, and this process is challenging [9, 10]. Lung ultrasound (LUS) is a rapid, bedside, goal-oriented, diagnostic test that is used to answer specific clinical questions, and its findings have been reported to be consistent with CT findings [11, 12]. Moreover, LUS can be quantitatively assessed, and the serial LUS score can be used to assess disease progression [13, 14]. However, there is a concern regarding whether the ultrasound findings can be reliably interpreted by examiners [15].

Therefore, in this study, we evaluated whether the serial LUS score could predict PMV and successful extubation in severe COVID-19 patients who require invasive mechanical ventilation and determined the inter-rater reliability (IRR) of the results among examiners.

Methods

Study design and population

This retrospective, single-centre, observational study included consecutive patients from Yokohama City University Medical Centre Advanced Critical Care and Emergency Centre, Japan, a hospital designated for treating severe COVID-19 patients, from 1 May 2020 to 28 February 2021. The following patients were included: those with a positive nasopharyngeal reverse transcription polymerase chain reaction for severe acute respiratory syndrome coronavirus 2; those aged > 18 years; and those who required mechanical ventilation over 48 h. The exclusion criteria were as follows: acute heart failure, interstitial pneumonia, other pulmonary diseases affecting image acquisition or suboptimal ultrasound window, missing ultrasound data, and patients' refusal to consent.

The study was approved by the institutional ethics board of Yokohama City University Medical Centre (approval number: B200200049). Written informed consent was waived, as ultrasound scanning of the lungs is considered a routine procedure.

Patient management

During invasive mechanical ventilation, sedation analgesia was managed at a Richmond Agitation Sedation Score under - 3 in patients with a strong respiratory effort, and muscle relaxant was also administered if necessary. The patient's respiratory effort was assessed based on the airway occlusion pressure (P0.1) and physical examination, and a P0.1 > 4 was considered a strong respiratory effort [16]. If the respiratory effort was calm, for example, a P0.1 \leq 4 and not using accessory respiratory muscles, daily spontaneous awakening trial (SAT) was performed, and the patient was managed according to the Pain, Agitation/Sedation, Delirium, Immobility, and Sleep Disruption guidelines [17].

Ventilation management was performed with pressure-controlled ventilation driving pressure < 14 and positive end-expiratory pressure (PEEP), based on a high PEEP table, from the acute respiratory distress syndrome (ARDS) net of respiratory frequency < 15 [18]. FiO₂ was adjusted to SpO₂ > 93%. The introduction criteria for prone ventilation according to the PROSEVA study were as follows: FiO₂ > 60% and PaO₂/ FiO₂ (P/F) ratio < 150 [19]. Prone ventilation was performed for at least 16 h. The introduction criteria for veno-venous ECMO according to the EOLIA trial were as follows: FiO₂ > 80% and P/F ratio < 80 for at least 6 h [20]

To evaluate extubation, spontaneous breathing trial was performed after SAT, and if the Rapid Shallow Breath Index was < 100, the patient was extubated [21]. Fluid management was attained by keeping the level of increase within 10% of the body weight. The final decision to extubate was made by a team that included the physician in charge. Tracheostomy was performed after 2 weeks of intubation owing to the risk of infection and according to the guideline [22].

Clinical data and outcomes

Data on patients' demographic characteristics, imaging and laboratory findings, comorbidities, complications, treatment for COVID-19, and outcomes were extracted from the electronic medical records. Laboratory tests and chest X-ray results were recorded every day after admission, for the assessment of COVID-19 progression and fibrosis. Similarly, LUS was performed daily, when possible (depending on the workload and available medical staff). CT was usually performed at the time of admission and thereafter, when the physician in charge deemed it necessary. The primary endpoint of the study was PMV, and the secondary endpoint was successful extubation. PMV was defined as the requirement of mechanical ventilation for > 21 days, including reintubation within 7 days, according to the National Association for Medical Direction of Respiratory Care Consensus Conference [1]. Successful extubation was defined as not requiring reintubation for > 7 days.

Performing LUS and chest X-ray scoring

LUS examinations were performed using an ultrasound equipment (GE Venue Go) with a 5-12-MHz linear transducer. LUS was performed at six points per hemithorax (superior and inferior regions anteriorly, laterally, and posteriorly), and bilaterally; a total of 12 regions were assessed with the probe placed at the intercostal space to obtain images widely. In each region, LUS signs, including B-lines/consolidation and pleural line abnormalities, were assessed, and the worst LUS signs were recorded according to a previous study [23].

B-lines/consolidations were quantitatively scored as follows: score 0, well-spaced B-lines < 3; score 1, well-spaced B-lines \geq 3; score 2, multiple coalescent B-lines; and score 3, lung consolidation. The pleural line was quantitatively scored as follows: score 0, normal; score 1, irregular pleural line; and score 2, blurred pleural line. The sum of both scores in all 12 zones yielded a final score (ranging from 0 to 60) that was defined as the LUS score.

The LUS score was evaluated by two emergency physicians who were experienced in performing LUS for over 25 cases and who were blinded to the clinical data [24]. The scores were evaluated independently, after which the final decisions were reached by consensus. Moreover, the LUS score was evaluated by five emergency physicians who were less experienced in performing LUS to assess IRR.

The radiographic assessment of the lung oedema (RALE) score was used to evaluate the chest X-ray [25]. To determine the RALE score, each radiograph was divided into quadrants, defined vertically by the vertebral column and horizontally by the first branch of the left main bronchus. Each quadrant was assigned a consolidation score of 0–4 to quantify the extent of the alveolar opacities based on the percentage of the quadrant with the opacification and a density score of 1–3, to quantify the overall density of the alveolar opacities, unless the consolidation score for that quadrant was 0. The density score (1 = hazy, 2 = moderate, and 3 = dense) allowed for more quantitative assessment of the density of opacification by quadrant. To calculate the final RALE score, the product of the consolidation and density score for each quadrant were summed for the final RALE score, ranging from 0 (no infiltrates) to 48 (dense consolidation in > 75% of each quadrant).

The RALE score was also evaluated by two experienced emergency physicians who were blinded to the clinical data. Scores were evaluated independently, after which the final decisions were reached by consensus.

IRR of the LUS score

The LUS examinations were anonymised before scoring. The raters were blinded to the clinical information and the assessments were performed by other raters. We assessed LUS score and the difference in the daily LUS

score changes (day 1 to 3 and day 1 to 7) between well-experienced and less-experienced emergency physicians. The changes in LUS scores were defined as follows: 0, no change from day 1 LUS score; 1, decreasing from day 1 LUS score; and 2, increasing from day 1 LUS score.

Statistical analysis

Continuous variables are expressed as the mean \pm standard deviation (SD) or median (interquartile range), as appropriate. Categorical variables are presented as frequencies (percentages). Analysis of variance (ANOVA) of continuous variables were evaluated using the Kruskal–Wallis test or two-way ANOVA owing to non-normally distributed data. Categorical variables were compared using the chi-square test or Fisher's exact test. To estimate the predictors of the outcomes, all potential predictors of the outcomes were included in the univariate analyses (Mann-Whitney U test). Variables with $P < 0.05$ in the univariate analysis were used in the receiver operator curve (ROC) analysis. The ROC analysis was performed to examine the sensitivity and specificity of prognosis parameters of the outcomes and determine the area under the curve (AUC) with the 95% confidence interval (CI). The best cut-off value, which was determined by the maximum Youden index, for predicting future events, was also determined.

The IRR of the LUS score was analysed using the kappa test [26]. The kappa value was interpreted using a scale interpretation by Altman's Benchmark Scale for the kappa [27]. The acceptable limit for the kappa value in this study was ≥ 0.41 , which was in moderate agreement.

All statistical analyses were performed using JMP®, Version 15. (SAS Institute Inc., Cary, NC.).

Results

Clinical characteristics

A total of 20 COVID-19 patients who met the inclusion criteria were identified. Six patients were excluded owing to missing ultrasound data. Table 1 summarises the baseline clinical characteristics of the patients with ($n = 11$) or without ($n = 9$) PMV. The patients comprised 16 men and four women, and their median age was 66 years (range, 56.3–73.8 years). The median duration between the occurrence of symptoms and hospital admission was 7 (range, 4–10) days. There was no significant difference in age, gender, body mass index, time from first symptom, pre-existing condition, Acute Physiologic Assessment and Chronic Health Evaluation II score, treatment, complications, outcome, and fluid balance. In imaging and laboratory findings, there were significant differences in the LUS score on days 3 and 7; XP score on day 7; and P/F ratio on day 7 ($p < 0.05$). There was no difference with fibrosis on CT at admission.

Table 1
Baseline clinical characteristics of the patients with or without PMV

Variable [frequency (%)/median (IQR)]	PMV(n = 11)		Non-PMV (n = 9)		p-value
Age (years)	66	[56–74]	65	[55.5–71]	0.595
Male sex	10	[50]	6	30]	0.285
BMI	28.4	[24.9–31.8]	25	[22.7–28.3]	0.183
Time from symptom	9	[5.5–10]	6	[3–7]	0.098
<i>Pre-existing condition</i>					
Hypertension	5	[25]	7	[35]	0.196
Diabetes	6	[30]	4	[20]	0.650
Chronic cardiac failure	1	[5]	1	[5]	0.889
Renal insufficiency	0	[0]	0	[0]	-
Malignancy	0	[0]	1	[5]	0.257
APACHE II score at the time of ICU admission	20	[16–22]	19	[15–24]	0.760
<i>Treatment</i>					
Antiviral therapy, n (%)	9	[45]	9	[45]	0.177
Glucocorticoid therapy, n (%)	9	[45]	9	[45]	0.177
Prone ventilation, n (%)	6	[30]	4	[20]	0.653
Extra-corporeal membrane oxygenation, n (%)	4	[20]	1	[5]	0.195
<i>Complications</i>					
Ventilator-associated pneumonia, n (%)	4	[20]	0	[0]	0.094
Septic shock, n (%)	0	[0]	0	[0]	-
Acute kidney injury, n (%)	0	[0]	0	[0]	-
Acute heart injury, n (%)	1	[5]	0	[0]	0.861
<i>Outcomes</i>					
Discharge, n (%)	10	{50]	9	[45]	0.353
Death, n (%)	1	[5]	0	[0]	0.353
<i>Fluid balance</i>					
Total fluid balance on day 3 (mL)	2734	[1414–4024]	2672	[2125–3861]	0.621

PMV, prolonged mechanical ventilation; IQR, interquartile range; BMI, body mass index; APACHE II, Acute Physiologic Assessment and Chronic Health Evaluation II; ICU, intensive care unit; CT, computed tomography; EF, ejection fraction; E/e', the ratio between early mitral inflow velocity and mitral annular early diastolic velocity; LUS, lung ultrasound; XP;X-ray; P/F ratio, PaO2/FiO2 ratio; CRP, C-reactive protein.

Variable [frequency (%)/median (IQR)]	PMV(n = 11)		Non-PMV (n = 9)		p-value
day 7	3934	[2205–4518]	1949	[371–3543]	0.063
<i>Imaging findings</i>					
Fibrosis on CT at admission, n (%)	1	[5]	0	[0]	0.354
Echocardiography (EF) on admission (%)	60	[50.0-63.1]	60	[55.8–66.8]	0.717
Echocardiography (E/e') on admission	12.4	[9.86–14.2]	8.75	[6.35–13.3]	0.387
LUS score on day 1	36	[35–36]	35	[31–36]	0.101
day 3	36	[36–39]	32	[25.5–36]	0.004
day 7	38	[36–41]	27	[23.5–30]	< 0.001
XP score on day 1	20	[19–30]	22	[19–23]	0.842
day 3	20	[16–30]	18	[9–24]	0.252
day 7	20	[16–30]	16	[9.5–17]	0.045
<i>Laboratory findings.</i>					
Lymphocytes on day 1 (%)	11	[6.95–13.3]	8.1	[4.8–12.3]	0.452
day 3	4.5	[3.2–6.95]	4.4	[3-8.7]	0.873
day 7	5.9	[4.35–10.6]	4.95	[3.68–13.3]	0.592
CRP on day 1 (mg/L)	13.4	[7.04–16.5]	16.1	[12.7–22.3]	0.239
day 3	6.98	[5.27–15.8]	7.4	[3.93–10.3]	0.500
day 7	5.87	[4.48–9.28]	3.83	[2.66–7.10]	0.102
D-dimer on day 1 (µg/L)	3.65	[1.58-38]	1.8	[1.4–2.7]	0.135
day 3	3.6	[2.15–4.4]	3.2	[2.6–5.4]	0.682
day 7	4.1	[2.3–7.5]	2.4	[1.9–5.5]	0.482
KL-6 on day 1 (U/mL)	405	[287–765]	278	[214–340]	0.094
day 3	753	[398–1049]	472	[352–1321]	0.558
day 7	787	[444–842]	509	[365–1016]	0.366
P/F ratio on day 1	220	[108–233]	144	[116–225]	0.381
day 3	220	[161–266]	250	[181–298]	0.381
day 7	225	[170–266]	300	[245–378]	0.023
PMV, prolonged mechanical ventilation; IQR, interquartile range; BMI, body mass index; APACHE II, Acute Physiologic Assessment and Chronic Health Evaluation II; ICU, intensive care unit; CT, computed tomography; EF, ejection fraction; E/e', the ratio between early mitral inflow velocity and mitral annular early diastolic velocity; LUS, lung ultrasound; XP, X-ray; P/F ratio, PaO ₂ /FiO ₂ ratio; CRP, C-reactive protein.					

Table 2 summarises the baseline clinical characteristics of the patients based on whether they had successful extubation or not. There was no significant difference in clinical characteristics between the groups. In imaging and laboratory findings, there were significant differences in the LUS score on day 3 and day 7; CRP levels on day 7; and P/F ratio on day 7 ($p < 0.05$).

Table 2
Baseline clinical characteristics of patients according to successful or non-successful extubation

Variable [frequency (%)/median (IQR)]	Successful extubation (n = 8)		Non-successful extubation (n = 12)		p-value
Age (years)	64.5	[54.3–68.5]	69	[56.3–74]	0.376
Male sex	5	[25]	11	[55]	0.255
BMI	25.3	[24.3–28.3]	26.9	[24–31.4]	0.537
Time from symptom	9	[4.8–10]	6	[3.3–9.3]	0.349
<i>Pre-existing condition</i>					
Hypertension	6	[30]	6	[30]	0.373
Diabetes	3	[15]	7	[35]	0.649
Chronic cardiac failure		[5]	1	[5]	0.889
Renal insufficiency	0	[0]	0	[0]	-
Malignancy	0	[0]	1	[5]	0.402
APACHE II score at the time of ICU admission	17.5	[15–23]	20	[17–23]	0.756
<i>Treatment</i>					
Antiviral therapy, n (%)	8	[40]	10	[50]	0.497
Glucocorticoid therapy, n (%)	8	[40]	10	[50]	0.497
Prone ventilation, n (%)	3	[15]	7	[35]	0.649
Extra-corporeal membrane oxygenation, n (%)	1	[5]	4	[20]	0.603
<i>Complications</i>					
Ventilator-associated pneumonia, n (%)	0	[0]	4	[20]	0.117
Septic shock, n (%)	0	[0]	0	[0]	-
Acute kidney injury, n (%)	0	[0]	0	[0]	-
Acute heart injury, n (%)	1	[5]	0	[0]	0.861
<i>Outcomes</i>					
Discharge, n (%)	8	[40]	11	[55]	0.304

IQR, interquartile range; BMI, body mass index; APACHE II, Acute Physiologic Assessment and Chronic Health Evaluation II; ICU, intensive care unit; CT, computed tomography; EF, ejection fraction; E/e', the ratio between early mitral inflow velocity and mitral annular early diastolic velocity; LUS, lung ultrasound; XP, X-ray; CRP, C-reactive protein.

Variable [frequency (%)/median (IQR)]	Successful extubation (n = 8)		Non-successful extubation (n = 12)		p-value
Death, n (%)	0	[0]	1	[5]	0.304
<i>Fluid balance</i>					
Total fluid balance on day 3 (mL)	2933	[2016–3136]	2754	[1474–4148]	0.938
day 7	2016	[370–3682]	2130	[1474–4485]	0.063
<i>Imaging findings</i>					
Fibrosis on CT at admission, n (%)	0	[0]	1	[5]	0.402
Echocardiography (EF) on admission (%)	61.4	[50-66.7]	54.4	[50-63.1]	0.713
Echocardiography (E/e') on admission	9.48	[6.35–13.3]	12.2	[9.86–14.2]	0.385
LUS score on day 1	34	[32–36]	34	[35–36]	0.371
day 3	32	[30–36]	36	[3–39]	0.027
day 7	28	[26–31]	37	[36–41]	0.003
XP score on day 1	22	[22–24]	22	[17–29]	0.640
day 3	18	[10–24]	21	[15–27]	0.614
day 7	14	[10–18]	21	[14–30]	0.147
<i>Laboratory findings.</i>					
Lymphocytes on day 1 (%)	7.44	[4.7–11.9]	11.3	[8.48–13.3]	0.328
day 3	3.75	[2.9–6.88]	4.75	[3.2–8.4]	0.447
day 7	4.3	[3.55–12.8]	7.3	[4.63–11.5]	0.297
CRP on day 1 (mg/L)	15.9	[12.7–21.1]	13.8	[7.81–21.2]	0.488
day 3	6.68	[3.76–9.14]	7.25	[5.47–23.2]	0.247
day 7	3.83	[2.64–4.79]	6.54	[4.48–12.3]	0.017
D-dimer on day 1 (µg/L)	3.65	[1.58-38]	1.8	[1.4–2.7]	0.135
day 3	7.44	[4.7–11.9]	3.2	[2.6–5.4]	0.682

IQR, interquartile range; BMI, body mass index; APACHE II, Acute Physiologic Assessment and Chronic Health Evaluation II; ICU, intensive care unit; CT, computed tomography; EF, ejection fraction; E/e', the ratio between early mitral inflow velocity and mitral annular early diastolic velocity; LUS, lung ultrasound; XP, X-ray; CRP, C-reactive protein.

Variable [frequency (%)/median (IQR)]	Successful extubation (n = 8)		Non-successful extubation (n = 12)		p-value
day 7	3.75	[2.9–6.88]	2.4	[1.9–5.5]	0.482
KL-6 on day 1 (U/mL)	282	[212–560]	382	[280–750]	0.160
day 3	592	[419–1625]	692	[355–1029]	0.758
day 7	545	[394–1267]	705	[397–835]	0.673
P/F ratio on day 1	148	[125–228]	210	[97.8–232]	0.754
day 3	230	[15–295]	225	[163–287]	0.643
day 7	302	[242–379]	227	[177–270]	0.034

IQR, interquartile range; BMI, body mass index; APACHE II, Acute Physiologic Assessment and Chronic Health Evaluation II; ICU, intensive care unit; CT, computed tomography; EF, ejection fraction; E/e', the ratio between early mitral inflow velocity and mitral annular early diastolic velocity; LUS, lung ultrasound; XP, X-ray; CRP, C-reactive protein.

ROC analysis

We conducted ROC curve analysis to assess the predictive values of LUS scores for days 3 and 7; XP score on day 7; P/F ratio on day 7 for PMV and LUS scores for days 3 and 7; CRP levels on day 7; P/F ratio day 7 for successful extubation.

Our analysis of the predictors showed that the AUCs of the LUS score on day 3; LUS score on day 7; XP score on day 7; and P/F ratio on day 7 were 0.88 (95% CI: 0.66–0.96); 0.98 (95% CI: 0.83–1.0); 0.77 (95% CI: 0.50–0.91); and 0.80 (95% CI: 0.54–0.93), respectively, for PMV ($p < 0.05$). A cut-off value of 35 for the LUS score on day 7 had a sensitivity of 100% and a specificity of 88.9% for the prediction of PMV (Fig. 1a).

For successful extubation, the AUCs of the LUS score on day 3; LUS score on day 7; CRP levels on day 7; and P/F ratio on day 7 were 0.79 (95% CI: 0.53–0.93); 0.90 (95% CI: 0.59–0.98); 0.82 (95% CI: 0.53–0.95); and 0.79 (95% CI: 0.52–0.93), respectively ($p < 0.05$). A cut-off value of 31 for the LUS score on day 7 had a sensitivity of 87.5% and a specificity of 91.7% for the prediction of successful extubation (Fig. 1b).

Analysis of the serial LUS score

Figure 2a shows the variations in LUS scores on days 1, 3, and 7 between the PMV and non-PMV groups. The serial LUS scores exhibited significant differences between the PMV and non-PMV groups ($p < 0.001$, two-way ANOVA). The serial LUS score on day 7 was significantly higher than that on day 1 in the PMV group ($p < 0.05$, Kruskal–Wallis test) but was significantly lower in the non-PMV group ($p < 0.05$, Kruskal–Wallis test). Moreover, the variations in LUS scores on days 1, 3, and 7 between the successful and non-successful extubation groups exhibited significant differences ($p = 0.001$, two-way ANOVA; Fig. 2b). The serial LUS score on day 7 was significantly lower than that of day 1 in the successful extubation group ($p < 0.05$, Kruskal–Wallis test) but was significantly higher in the non-successful extubation group ($p < 0.05$, Kruskal–Wallis test). We showed examples of PMV and successful extubation cases that were monitored using ultrasound and CT (Fig. 3) and found that

higher LUS scores on day 7 predicted PMV, while lower LUS scores on day 7 predicted successful extubation in patients with severe COVID-19.

IRR

There was no IRR agreement in the LUS score between well-experienced and less-experienced emergency physician examiners ($\kappa = 0.01$ [95% CI: 0-0.04]). There was fair IRR agreement in the LUS score changes on day 1 to 3 ($\kappa = 0.23$ [95% CI: 0.08–0.38]) and slight IRR agreement in the LUS score changes on day 1 to 7 ($\kappa = 0.15$ [95% CI: 0-0.31]). The results are shown in an additional table file (see Additional files 1 and 2).

Discussion

Several studies on the use of LUS in COVID-19 patients, which use CT as the reference standard, have indicated that LUS on admission may predict mortality or the need for invasive mechanical ventilation [13, 14, 23]. However, few studies that assessed the serial LUS scores could predict the prognosis of lung injury. In this study, we showed that higher LUS score on day 7 was a predictor for PMV, while lower LUS score on day 7 was a predictor for successful extubation in patients with severe COVID-19.

If patients receive PMV, they are usually excluded as candidates for ECMO, and with limited resources during a pandemic, this may be considered a withdrawal of treatment [4, 5]. The rapid surge of medical needs depletes the ventilators and ICU beds, making the use of anaesthetic machines instead of ventilators compulsory [6]. Therefore, it is very important to predict whether patients will require PMV or can be extubated if ventilatory management becomes necessary. If we can predict the need for PMV early, we can consider transferring the patient to an ECMO centre before ECMO is no longer applicable. Furthermore, the ability to predict PMV allows for the appropriate allocation of medical resources, including ICU beds.

Gattinoni et al. reported variations in the respiratory mechanics profiles of invasively ventilated patients with COVID-19 pneumonitis, and the following two clinical phenotypes were identified: (1) type L, which is characterised by low elastance, a low ventilation-to-perfusion ratio, a low lung weight, and a low recruitability, and (2) type H, which is characterised by high elastance, a pronounced right-to-left shunt, a high lung weight, and a high recruitability [28]. The transition from Type L to Type H may be because of the worsening of COVID-19 severity, or an injury caused by high-stress ventilation and patient self-inflicted ventilation (P-SILI) [29, 30]. The depth of the negative intrathoracic pressure may also play a key role in the phenotype shift. If P-SILI is a concern in COVID-19 patients, early intubation is recommended, and adequate sedation and analgesia should be administered to suppress spontaneous breathing [30, 31]. However, the patient's condition should be evaluated to determine how long the lungs should be rested and when the lungs should be used. Excessive sedation and analgesia may result in unsuccessful extubation, which is a risk factor for PMV [32].

Follow-up CT in ARDS patients, including COVID patients, could demonstrate the progression of lung pathology [9–11, 33]. Pulmonary fibroproliferation, assessed using CT, in patients with ARDS, which is induced by COVID-19, predicts increased mortality and increased susceptibility to multiple organ failure, including ventilator dependency and its associated outcomes [8, 9, 34]. However, in a pandemic, the transportation of critically ill ventilated patients to radiology facilities is challenging, especially for ECMO-managed patients [11, 12]. LUS is a fast, non-invasive, sensitive, and quantitative tool to assess multiple pulmonary pathologies, such as pulmonary oedema, pneumonia, and interstitial lung disease [35]. Recently, it has been shown that LUS findings are similar

to chest CT findings, and they are superior to chest radiography findings for evaluating patients with COVID-19 [36, 37]. Indeed, we reported the usefulness of LUS as the sole imaging modality with bedside accessibility to patients for the timely identification of pulmonary condition, thus reducing the risk of moving unstable ECMO patients [14]. Furthermore, de Almeida Monteiro et al. showed a histological background that supports the fact that LUS can be used to characterise the progression and severity of lung damage in severe COVID-19 [38]. Therefore, LUS may have very useful imaging findings in COVID-19 patients, which are consistent with CT and pathologic findings.

In this study, there was no difference in water balance or cardiac function according to the outcome. Therefore, we believe that the worsening of LUS scores can be used to evaluate lung injury, such as fibrosis. Although sialylated carbohydrate antigen KL-6 (KL-6) is usually used as a biomarker to evaluate lung fibrosis and can predict severity in COVID-19 patients [39], there was no significant difference between PMV and non-PMV in our study. Compared with a previous report [39], our study included only severely ill patients, and KL-6 was indeed high, suggesting the severity of the disease but not prognosis.

There are few reports of COVID-19 patients who met the usual extubation criteria but were subsequently reintubated [13, 40]. Moreover, CT at the time of reintubation shows progressive fibrosis of the lung [13, 40]. In our study, three patients in the PMV group met our extubation criteria and were once extubated but were reintubated within 7 days. The reason for the reintubation could be exacerbation of the respiratory workload owing to the lung fibrosis. It has been reported that the success rate of extubation is higher when respiratory effort and diaphragmatic muscle strength are added to the evaluation, besides the conventional extubation criteria [41, 42]. Based on our results and previous reports of ultrasound evaluation of the diaphragm [41], we believe that ultrasound assessment may be considered in future extubation criteria.

Another finding of our study was that although LUS was considered useful, the IRR agreement was low. Previous studies have reported that high IRR of B-line is suggestive of pulmonary oedema, while low IRR is suggestive of pleural thickness and abnormalities observed in ARDS, including COVID-19 and lung fibrosis [15, 43]. The results of our study are similar with the results of previous studies [15]. Although LUS is a rapid, bedside, goal-oriented, diagnostic test, the IRR variation is a common problem in point of care ultrasound. To overcome this problem, we believe it is very important to perform a study on the automatic ultrasound judgement such as deep learning [44, 45].

This study had some limitations. This was a single-centre study with a relatively limited sample size; this could limit the generalisability of our results. Therefore, further multicentre studies with a larger sample size are needed to assess our findings. Secondly, it is suggested that respiratory muscle strength, including diaphragmatic functions, affect PMV and successful extubation; however, this was not assessed in this study. Finally, a daily comparison between the LUS score and chest CT was not performed because we had extremely limited CT imaging data (almost only available on admission).

Conclusions

In this study, we showed that higher LUS score on day 7 was a predictor for PMV, while lower LUS score on day 7 was a predictor for successful extubation in patients with severe COVID-19. However, although LUS was useful, the IRR was low. Therefore, we should consider automatic ultrasound judgement, such as deep learning, to overcome IRR disagreement.

Abbreviations

ARDS, acute respiratory distress syndrome

AUC, area under the curve

CI, confidence interval

CNN, Convolutional Neural Network

COVID-19, coronavirus disease

CT, computed tomography

ECMO, extra-corporeal membranous oxygenation

ICU, intensive care unit

IRR, inter-rater reliability

LUS, lung ultrasound

P.01, airway occlusion pressure

PEEP, positive end-expiratory pressure

PMV, prolonged mechanical ventilation

P-SILI, patient self-inflicted ventilation

RALE, radiographic assessment of lung oedema

ROC, receiver operator curve

SAT, spontaneous awakening trial

Declarations

Ethics approval and consent to participate

The study was approved by the institutional ethics board of Yokohama City University Medical Centre (approval number: B200200049). Written informed consent was waived, as ultrasound scanning of the lungs is considered a routine procedure.

Consent for publication

Consent for publication was obtained from all the participants. Written informed consent was waived, as ultrasound scanning of the lungs is considered a routine procedure.

Data availability statements

All data generated or analysed during this study are included in this published article.

Competing interest

The authors declare that they have no competing interests.

Funding

We did not receive any funding support for this study.

Authors' contributions

TA and IT designed the study and are guarantors of the paper. TA and MI prepared the draft and finalised the manuscript. AO and TA helped with data analysis and interpretation of the results. AO and HY were involved in clinical data collection and arrangement. All authors read and approved the final manuscript.

Acknowledgements

We would like to thank all the staff of Yokohama City University Medical Centre, Advanced Critical Care Centre, involved in this study.

References

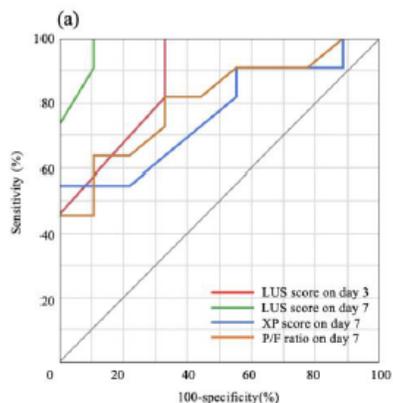
1. MacIntyre NR, Epstein SK, Carson S, Scheinhorn D, Christopher K, Muldoon S, et al. Management of patients requiring prolonged mechanical ventilation: report of a NAMDRC consensus conference. *Chest*. 2005;128:3937-54.
2. Serafim RB, Póvoa P, Souza-Dantas V, Kalil AC, Salluh JIF. Clinical course and outcomes of critically ill patients with COVID-19 infection: a systematic review. *Clin Microbiol Infect*. 2021;27:47–54.
3. Argenziano MG, Bruce SL, Slater CL, Tiao JR, Baldwin MR, Barr RG, et al. Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ*. 2020;369:m1996.
4. Wunsch H. Mechanical ventilation in COVID-19: Interpreting the current epidemiology. *Am J Respir Crit Care Med*. 2020;202:1–4.
5. Extracorporeal Life Support Organization (ELSO). Guidelines for adult respiratory failure. August, 2017. https://www.else.org/Portals/0/ELSO%20Guidelines%20For%20Adult%20Respiratory%20Failure%201_4.pdf. Accessed 10 June 2021.
6. 6.: Austin PN, Branson RD. Anesthesia machine ventilators for use in critical care during the COVID-19 pandemic [published online ahead of print]. *Respir Care*. 2021;respcare.08799.
7. Takhar A, Surda P, Ahmad I, Amin N, Arora A, Camporota L, et al. Timing of tracheostomy for prolonged respiratory wean in critically ill coronavirus disease 2019 patients: a machine learning approach. *Crit Care Explor*. 2020;2:e0279.
8. Fang Y, Zhou J, Ding X, Ling G, Yu S. Pulmonary fibrosis in critical ill patients recovered from COVID-19 pneumonia: preliminary experience. *Am J Emerg Med*. 2020;38:2134–8.

9. Chen X, Tang Y, Mo Y, Li S, Lin D, Yang Z, et al. A diagnostic model for coronavirus disease 2019 (COVID-19) based on radiological semantic and clinical features: a multi-center study. *Eur Radiol.* 2020;30:4893–902.
10. Pan F, Ye T, Sun P, Gui S, Liang B, Li L, et al. Time course of lung changes at chest CT during recovery from coronavirus Disease 2019 (COVID-19). *Radiology.* 2020;295:715–21.
11. Hussain A, Via G, Melniker L, Goffi A, Tavazzi G, Neri L, et al. Multi-organ point-of-care ultrasound for COVID-19 (PoCUS4COVID): international expert consensus. *Crit Care.* 2020;24:702.
12. Peng QY, Wang XT, Zhang LN, Chinese Critical Care Ultrasound Study Group (CCUSG). Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic. *Intensive Care Med.* 2020;46:849–50.
13. Dargent A, Chatelain E, Kreitmann L, Quenot J, Cour M, Argaud L, et al. Lung ultrasound score to monitor COVID-19 pneumonia progression in patients with ARDS. *PLoS One.* 2020;15:e0236312.
14. Taniguchi H, Ohta S, Honzawa H, Takahashi K, Iwashita M, Abe T, et al. Usefulness of serial lung ultrasound for a severe COVID-19 patient on extracorporeal membrane oxygenation. *Respir Med Case Rep.* 2021;33:101383.
15. Kumar A, Weng Y, Graglia S, Chung S, Duanmu Y, Lalani F, et al. Interobserver agreement of lung ultrasound findings of COVID-19 [published online ahead of print]. *J Ultrasound Med.* 2021;10.1002/jum.15620.
16. Telias I, Junhasavasdikul D, Rittayamai N, Piquilloud L, Chen L, Ferguson ND, et al. Airway occlusion pressure as an estimate of respiratory drive and inspiratory effort during assisted ventilation. *Am J Respir Crit Care Med.* 2020;201:1086–98.
17. Devlin JW, Skrobik Y, Gélinas C, Needham DM, Slooter AJC, Pandharipande PP, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult patients in the ICU. *Crit Care Med.* 2018;46:e825-73.
18. NIH-NHLBI ARDS Clinical Network. Mechanical ventilation protocol summary. http://www.ardsnet.org/files/ventilator_protocol_2008-07.pdf. Accessed 10 June 2021.
19. Guérin C, Reignier J, Richard JC, Beuret P, Gacouin A, Boulain T, et al. Prone positioning in severe acute respiratory distress syndrome. *N Engl J Med.* 2013;368:2159–68.
20. Combes A, Hajage D, Capellier G, Demoule A, Lavoué S, Guervilly C, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. *N Engl J Med.* 2018;378:1965–75.
21. Baptistella AR, Sarmiento FJ, da Silva KR, Baptistella SF, Taglietti M, Zuquello R, et al. Predictive factors of weaning from mechanical ventilation and extubation outcome: a systematic review. *J Crit Care.* 2018;48:56–62.
22. Takhar A, Walker A, Tricklebank S, Wyncoll D, Hart N, Jacob T, et al. Recommendation of a practical guideline for safe tracheostomy during the COVID-19 pandemic. *Eur Arch Otorhinolaryngol.* 2020;277:2173–84.
23. Ji L, Cao C, Gao Y, Zhang W, Xie Y, Duan Y, et al. Prognostic value of bedside lung ultrasound score in patients with COVID-19. *Crit Care.* 2020;24:700.
24. Rouby JJ, Arbelot C, Gao Y, Zhang M, Lv J, An Y, et al. Training for lung ultrasound score measurement in critically ill patients. *Am J Respir Crit Care Med.* 2018;198:398–401.
25. Warren MA, Zhao Z, Koyama T, Bastarache JA, Shaver CM, Semler MW, et al. Severity scoring of lung oedema on the chest radiograph is associated with clinical outcomes in ARDS. *Thorax.* 2018;73:840–6.
26. Agresti A. Categorical data analysis. New York: John Wiley & Sons; 1990.

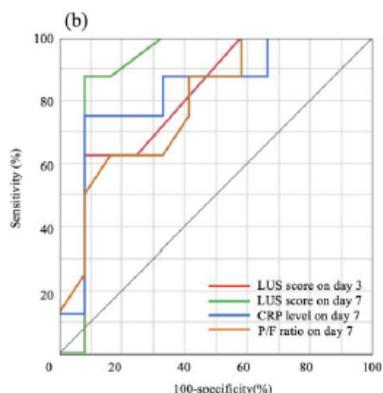
27. McCray G. Assessing inter-rater agreement for nominal judgement variables. Paper presented at the Language Testing Forum. Nottingham, November 15–17; 2013.
28. Gattinoni L, Chiumello D, Caironi P, Busana M, Romitti F, Brazzi L, et al. COVID-19 pneumonia: different respiratory treatments for different phenotypes? *Intensive Care Med.* 2020;46:1099–102.
29. Gattinoni L, Coppola S, Cressoni M, Busana M, Rossi S, Chiumello D. COVID-19 does not lead to a "typical" acute respiratory distress syndrome. *Am J Respir Crit Care Med.* 2020;201:1299–300.
30. Brochard L, Slutsky A, Pesenti A. Mechanical ventilation to minimize progression of lung injury in acute respiratory failure. *Am J Respir Crit Care Med.* 2017;195:438–42.
31. Pearson SD, Patel BK. Evolving targets for sedation during mechanical ventilation. *Curr Opin Crit Care.* 2020;26:47–52.
32. Rose L, McGinlay M, Amin R, Burns KE, Connolly B, Hart N, et al. Variation in definition of prolonged mechanical ventilation. *Respir Care.* 2017;62:1324–32.
33. Ichikado K, Muranaka H, Gushima Y, Kotani T, Nader HM, Fujimoto K, et al. Fibroproliferative changes on high-resolution CT in the acute respiratory distress syndrome predict mortality and ventilator dependency: a prospective observational cohort study. *BMJ Open.* 2012;2:e000545.
34. Zeng Z, Xiang M, Guan H, Liu Y, Zhang H, Xia L, et al. Early fibroproliferative signs on high-resolution CT are associated with mortality in COVID-19 pneumonia patients with ARDS: a retrospective study. *Ther Adv Chronic Dis.* 2021;12:2040622320982171.
35. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, et al. International evidence-based recommendations for point-of-care lung ultrasound. *Intensive Care Med.* 2012;38:577–91.
36. Lu W, Zhang S, Chen B, Chen J, Xian J, Lin Y, et al. A clinical study of noninvasive assessment of lung lesions in patients with coronavirus disease-19 (COVID-19) by bedside ultrasound. *Ultraschall Med.* 2020;41:300–07.
37. Yang Y, Huang Y, Gao F, Yuan L, Wang Z. Lung ultrasonography versus chest CT in COVID-19 pneumonia: a two-centered retrospective comparison study from China. *Intensive Care Med.* 2020;46:1761–3.
38. de Almeida Monteiro RA, Duarte-Neto AN, Ferraz da Silva LF, Pierre de Oliveira E, Toledo do Nascimento EC, Mauad T, et al. Ultrasound assessment of pulmonary fibroproliferative changes in severe COVID-19: a quantitative correlation study with histopathological findings. *Intensive Care Med.* 2021;47:199–207.
39. d'Alessandro M, Cameli P, Refini RM, Bergantini L, Alonzi V, Lanzarone N, et al. Serum KL-6 concentrations as a novel biomarker of severe COVID-19. *J Med Virol.* 2020;92:2216–20.
40. Zhang J, He X, Hu J, Li T. Failure of early extubation among cases of coronavirus disease-19 respiratory failure: Case report and clinical experience. *Medicine.* 2020;99:e20843.
41. Goligher EC, Dres M, Fan E, Rubenfeld GD, Scales DC, Herridge MS, et al. Mechanical ventilation-induced diaphragm atrophy strongly impacts clinical outcomes. *Am J Respir Crit Care Med.* 2018;197:204–13.
42. Wawrzyniak IC, Rios Vieira SR, Victorino JA. Weaning from mechanical ventilation in ARDS: aspects to think about for better understanding, evaluation, and management. *Biomed Res Int.* 2018;2018:5423639.
43. Anderson KL, Fields JM, Panebianco NL, Jenq KY, Marin J, Dean AJ. Inter-rater reliability of quantifying pleural B-lines using multiple counting methods. *J Ultrasound Med.* 2013;32:115–20.
44. Roy S, Menapace W, Oei S, Luijten B, Fini E, Saltori C, et al. Deep learning for classification and localization of COVID-19 markers in point-of-care lung ultrasound. *IEEE Trans Med Imaging.* 2020;39:2676–87.

45. Arntfield R, VanBerlo B, Alaifan T, Phelps N, White M, Chaudhary R, et al. Development of a convolutional neural network to differentiate among the etiology of similar appearing pathological B lines on lung ultrasound: a deep learning study. *BMJ Open*. 2021;11:e045120.

Figures



Variable	AUC (95% CI)	P-value	Cut-off	Sensitivity	Specificity
LUS score on day 3	0.88 [0.66-0.96]	<.001	32	100%	66.7%
LUS score on day 7	0.98 [0.83-1.0]	<.001	35	100%	88.9%
XP score on day 7	0.77 [0.50-0.91]	0.012	20	54.5%	100%
P/F ratio on day 7	0.80 [0.54-0.93]	0.02	230	63.6	88.9%



Variable	AUC (95% CI)	P-value	Cut-off	Sensitivity	Specificity
LUS score on day 3	0.79 [0.53-0.93]	0.169	34	62.5%	91.7%
LUS score on day 7	0.90 [0.59-0.98]	0.005	31	87.5%	91.7%
CRP level on day 7	0.82 [0.53-0.95]	0.021	4.33	75%	91.7%
P/F ratio on day 7	0.79 [0.52-0.93]	0.02	300	62.5%	83.3%

Figure 1

Receiver operator curve analysis for predicting PMV and successful extubation PMV, prolonged mechanical ventilation; AUC, area under the curve; CI, confidence interval; LUS, lung ultrasound; XP score, X-ray score; P/F ratio, PaO₂/FiO₂ ratio; CRP, C-reactive protein

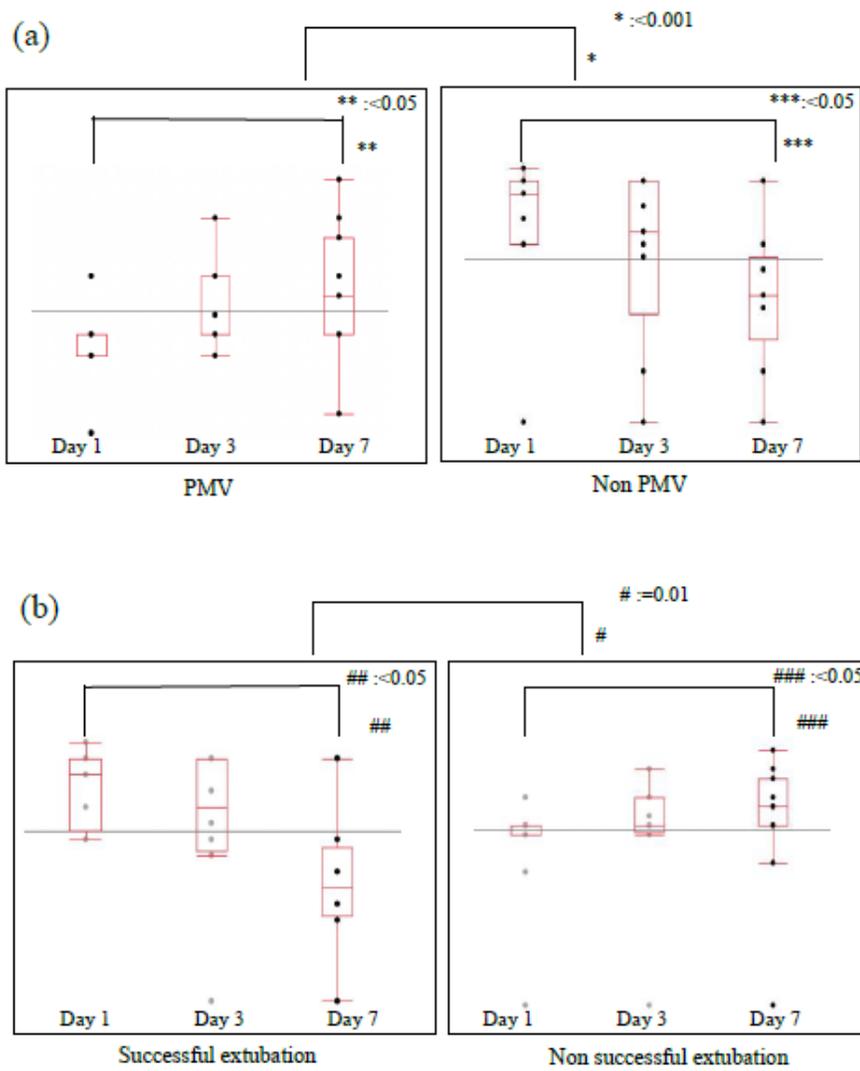


Figure 2

Analysis of variance of the LUS score PMV, prolonged mechanical ventilation; LUS, lung ultrasound

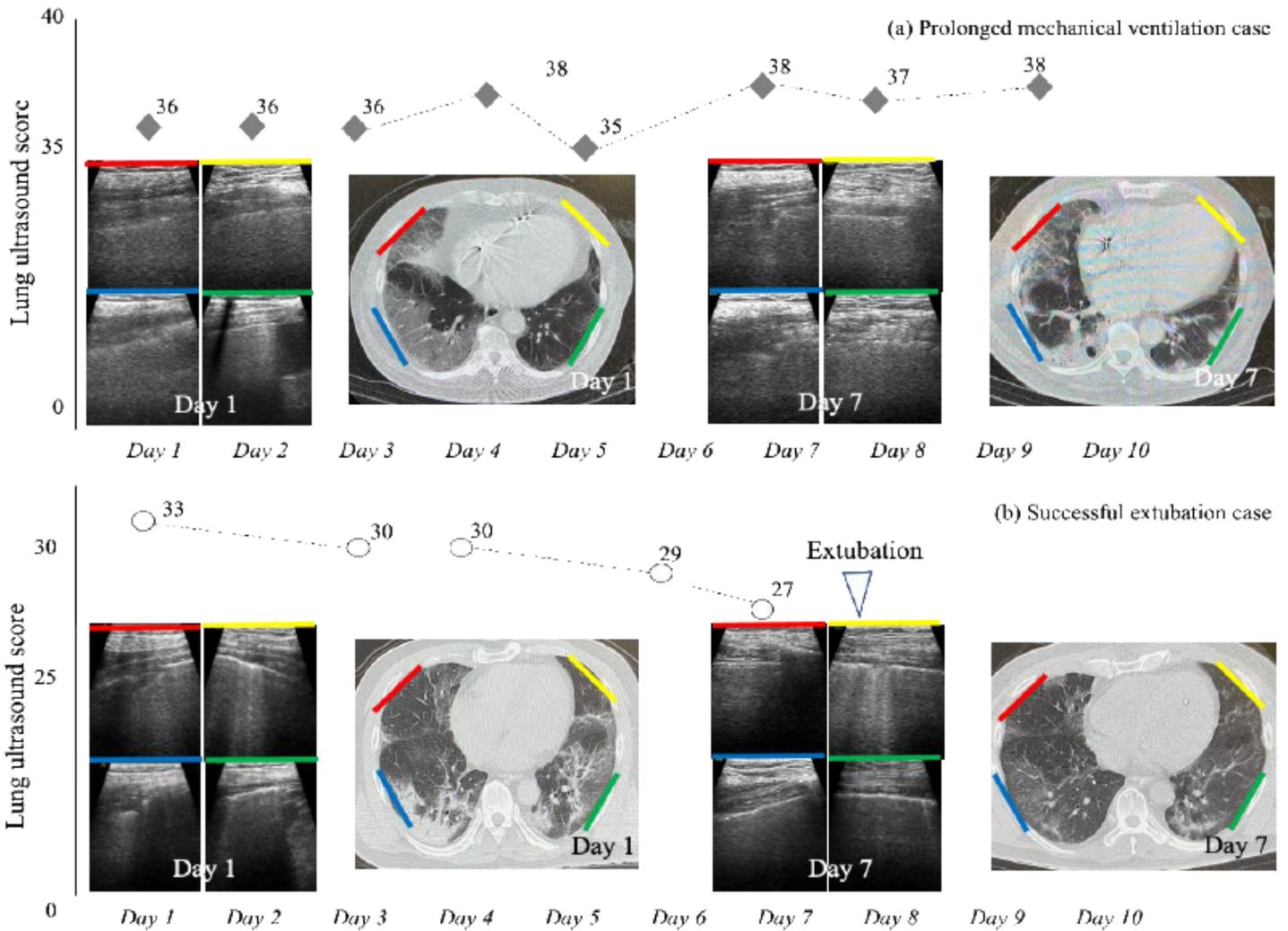


Figure 3

Example cases of PMV and successful extubation that could have been monitored using LUS and CT. PMV, prolonged mechanical ventilation; LUS, lung ultrasound; CT, computed tomography

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfilefinal.pdf](#)