On the Impact of Different Lung Ultrasound Imaging Protocols in the Evaluation of Patients Affected by Coronavirus Disease 2019

How Many Acquisitions Are Needed?

Federico Mento, MSc ⬤*, Tiziano Perrone, MD, Veronica Narvena Macioce, Francesco Tursi, MD,*
Danilo Buonsenso, MD ⬤*, Elena Torri, MD, Andrea Smargiassi, MD* ⬤*, Riccardo Inchingolo, MD* ⬤*,*
Gino Soldati, MD, Libertario Demi, PhD ⬤

Address correspondence to Libertario Demi, PhD Department of Information Engineering and Computer Science, University of Trento, via Sommarive 9, 38123, Trento, Italy.

E-mail: libertario.demi@unitn.it

Abbreviations
COVID-19, coronavirus disease 2019; LUS, lung ultrasound

Lung ultrasound (LUS) is currently being extensively used for the evaluation of patients affected by coronavirus disease 2019. In the past months, several imaging protocols have been proposed in the literature. However, how the different protocols would compare when applied to the same patients had not been investigated yet. To this end, in this multicenter study, we analyzed the outcomes of 4 different LUS imaging protocols, respectively based on 4, 8, 12, and 14 LUS acquisitions, on data from 88 patients. Results show how a 12-area acquisition system seems to be a good tradeoff between the acquisition time and accuracy.

Key Words—coronavirus disease 2019; lung ultrasound; point-of-care ultrasound; ultrasound imaging

The use of lung ultrasound (LUS) has widely spread during the coronavirus disease 2019 (COVID-19) pandemic.[1] Thanks to its main characteristics (real-time imaging, safety, portability, and availability), LUS can be used for stratification and monitoring of patients with COVID-19, as well as for triage of symptomatic patients. Lung ultrasound had shown its capability to detect alterations along the lung surface even before the COVID-19 pandemic. Nevertheless, being mainly based on the interpretation of imaging artifacts, eg, horizontal (A-lines) and vertical (B-lines) artifacts, LUS leads to qualitative and subjective analyses.[2] Standardization is therefore a crucial aspect, especially for defining the optimal imaging settings and the most suitable transducer. Moreover, it is fundamental to favor the reproducibility and increase the accuracy of LUS findings. In this pandemic context, a proposal for LUS standardization was recently presented, in which a 4-level score was defined, together with the acquisition time, transducers, imaging settings, and landmarks.[3] Other protocols and scanning procedures were proposed both during the COVID-19 pandemic and previously.[4] Finding the optimal tradeoff between the acquisition time and accuracy represents one of the main challenges. Specifically, to reduce the scanning time, a lower number of acquisition areas is preferred, but this could lead to underestimations.

In this multicenter study, we investigated how investigating different scanning areas influences the evaluation of a patient. Specifically, we acquired LUS videos from 88 patients by following the protocol of Soldati et al[3] based on 14 scanning areas per patient. A score ranging from 0 to 3 was assigned to each video, according to the 4-level scoring system.[3] Successively, each patient was classified according to the highest score (from 0 to 3) assigned to his or her 14 videos. Finally, we considered different subgroups of scanning areas to reevaluate the worst score of each patient and compared this value with the worst score obtained by the reference protocol.[3] The percentage of agreement was hence computed by summing the number of patients sharing the same worst score from the reference protocol[3] and dividing it by the total number of patients. As done by Smargiassi et al,[5] we first evaluated the level of agreement by separately considering only the anterior (labeled 11, 12, 13, and 14[3]), lateral (labeled 7, 8, 9, and 10[3]), and posterior (labeled 1, 2, 3, 4, 5, and 6[3]) areas. Then we computed the level of agreement for the 3 protocols based on 4, 8, and 12 scanning areas.[4] We named the analyzed systems system 1 (4 areas), system 2 (8 areas), system 3 (12 areas), and system 4 (14 areas). To further investigate the specific impact of the posterior areas, we analyzed the level of agreement for 3 modified versions of system 4 (with 10 areas instead of 14) obtained by considering all of the anterior and lateral areas together with the basal posteriors (1 and 4), middle posteriors (2 and 5), or apical posteriors (3 and 6).

The patient population consisted of 88 patients with a diagnosis of COVID-19 positivity by a reverse transcription polymerase chain reaction swab test. Of the 88 patients (48 male and 40 female), 9 were examined within the Lodi General Hospital (Lodi, Italy), 29 within the Fondazione Policlinico Universitario Agostino Gemelli (Rome, Italy), and 50 within the Fondazione Policlinico San Matteo (Pavia, Italy). This study was part of a protocol that has been registered (NCT04322487) and received approval from the Ethical Committee of the Fondazione Policlinico Universitario Agostino Gemelli, Istituto di Ricovero e Cura a Carattere Scientifico (protocol 0015884/20 ID 3117), of Milano area 1, the Azienda Socio-Sanitaria Territoriale Fatebenefratelli-Sacco (protocol N0031981), and the Fondazione Policlinico Universitario San Matteo (protocol 20200063198). All patients gave informed consent. The patients' ages ranged from 26 to 95 years (average, 59.9 years).

Figure 1 shows the distributions of scores for anterior, lateral, and posterior areas. The highest percentage of score 0 (35.23%) was observed in the anterior areas. In contrast, posterior areas showed the highest percentage of score 3 (30.87%) compared to anterior (10.80%) and lateral (16.76%) areas. Jointly considering score 2 and score 3, the overall percentages for the anterior, lateral, and posterior areas were 38.35%, 54.26%, and 58.52%, respectively. Therefore, the highest scores were focused on the lateral and posterior areas, whereas the lowest scores were focused on the anterior area. As a consequence, the levels of agreement with system 4 for just the anterior, lateral, and posterior areas were 40%, 58%, and 92% (Figure 1). These results were consistent with the results achieved by Smargiassi et al[5] and highlight the importance of the posterior areas.

Figure 2, top, shows how the distributions of scores vary with different systems: ie, system 1 (scanned areas, 7, 9, 12, and 14[3]), system 2 (scanned areas, 7–14[3]), system 3 (scanned areas, 1, 3, 4, 6, and 7–14[3]), and system 4 (scanned areas, 1–14[3]). Of particular interest is the distribution of score 3, whose percentage values were 14.78%, 14.20%, 19.90%, and 21.72% for systems 1, 2, 3, and 4, respectively. Consequently, the levels of agreement with system 4 for systems 1, 2, and 3 were 57%, 70%, 99% (Figure 2, top). This result suggests that the 12-area scanning protocol[4] leads to performance comparable with that of the reference protocol (14 areas).[3] Figure 2, bottom, shows the score distributions for the posterior areas. The highest percentage of score 0 (52.63%) was observed when scanning the apical posterior areas, whereas the highest percentage of score 3 (46.01%) was observed when scanning the basal posterior areas. Consequently, the levels of agreement between system 4 and the modified systems[4] were 80%, 89%, and 95% when the scanned posterior areas were the apical, middle, and basal, respectively. This highlights the importance of scanning basal posterior areas.

To conclude, the results show that the posterior areas are fundamental to capture the most important findings in patients with COVID-19. Moreover, the 12-area system seems to be a better tradeoff between the acquisition time and accuracy. However, the 10-area system also seems to be sufficiently accurate, as long as the basal posteriors are included.

Figure 1. The overall distribution of assigned scores divided per specific area (anterior, lateral, and posterior) is depicted on the left side. The percentage of scores assigned for each area and for each patient is depicted in the center. The level of agreement for all of the patients' scanning in the different areas is shown on the right side (For further details about the structure of agreement graphs, see Smargiassi et al.[5]) Each patient was classified according to the worst score. The reference system is system 4.
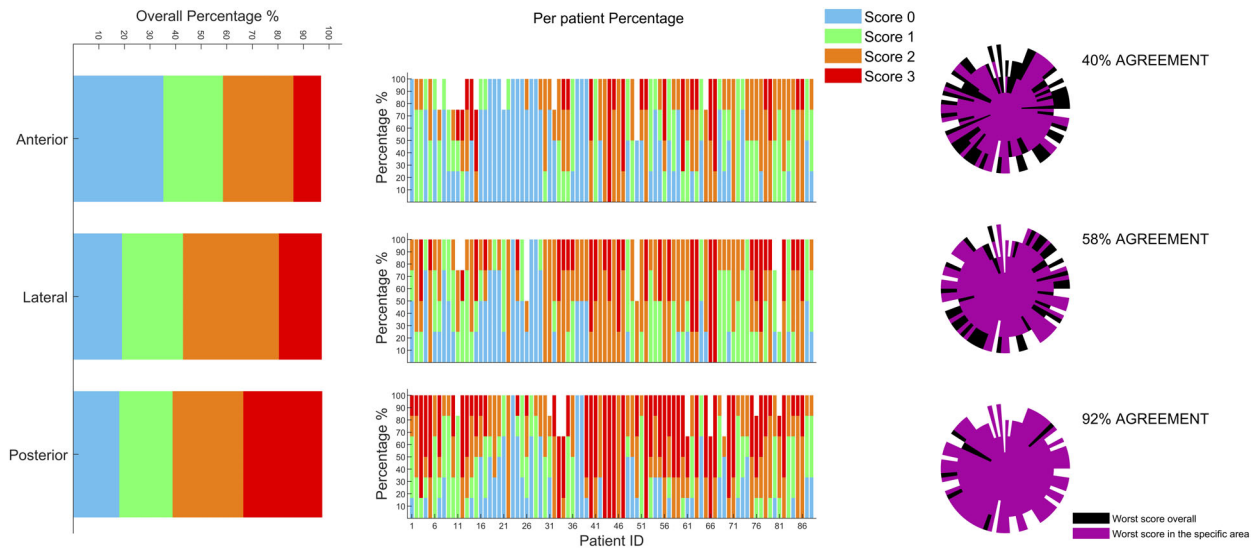

Figure 2. The overall distribution of score considering the 4 systems is depicted on the top left. The level of agreement between systems 1, 2, and 3 with respect to system 4 is also shown on the top. (For further details about the structure of agreement graphs, see Smargiassi et al.[5]) The distribution of each score in the posterior areas (basal, middle, and apical) is also depicted on the bottom left. The level of agreement between the 3 modified versions of system 4 (10 zones instead of 14: ie, all of the anterior and lateral areas together with apical posteriors, middle posteriors, or basal posteriors) with respect to system 4 is also shown on the top.
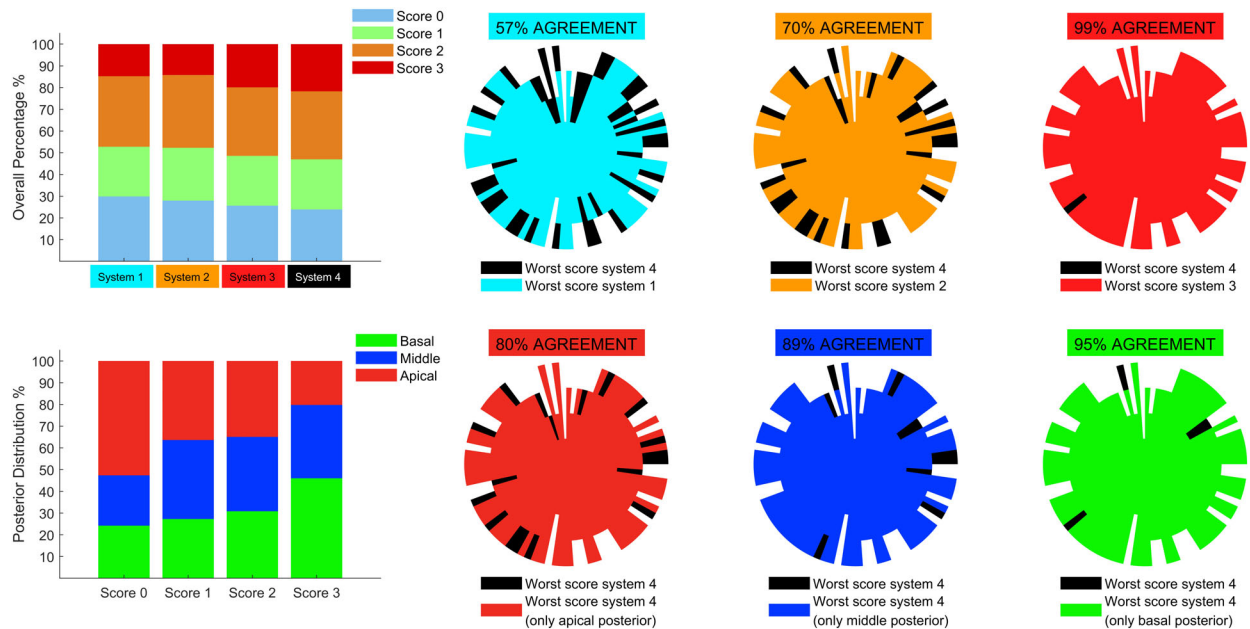
References

1. Soldati G, Smargiassi A, Inchingolo R, et al. Is there a role for lung ultrasound during the COVID-19 pandemic? *J Ultrasound Med* 2020; 39:1459–1462.

2. Demi L. Lung ultrasound: the future ahead and the lessons learned from COVID-19. *J Acoust Soc Am* 2020; 39:2146–2150.

3. Soldati G, Smargiassi A, Inchingolo R, et al. Proposal for international standardization of the use of lung ultrasound for patients with COVID-19. *J Ultrasound Med* 2020; 39: 1413–1419.

4. Allinovi M, Parise A, Giacalone M, et al. Lung ultrasound may support diagnosis and monitoring of COVID-19 pneumonia. *Ultrasound Med Biol* 2020; 46:2908–2917.

5. Smargiassi A, Soldati G, Torri E, et al. Lung ultrasound for COVID-19 patchy pneumonia: extended or limited evaluations [published online ahead of print August 20, 2020]? *J Ultrasound Med.* doi: https://doi.org/10.1002/jum.15428.