Background Lung ultrasound B-lines (LUS) identification is presently an evolving method for identification and follow up of interstitial lung water accumulation. Although it had been increasingly used, its limitation relates to the dependency of the reading skills of the operators, necessary to level their report that can impact on its diagnostic accuracy. Purpose We sought to determine the accuracy of an in-training group of medical residents (MDR) of various specialties that routinely use LUS everyday in bedside practice to determine presence and grading of pulmonary congestion. Methods Sixteen MDR in 4 University hospitals read a set of 20 video-clips of LUS of different patients. For each set they gave an answer of number of B-Lines varying from 0 (clear lung, only A-lines) to 10 (severe lung, coalescent B-lines). Diagnostic reference pattern (gold standard) was considered the concordant answer of two experts in this setting and very experienced readers. The answer was considered correct if equals to gold standard ±1 in e.g., gold-standard 5 B-lines; correct answer 4, 5 and 6. Sixty naive cardiologists qualified for B-lines reading working in 52 centers (16 countries) of a network of studies of stress echocardiography that also read the set of videos after taking an obligatory online course of 2-hours duration, also read the same set of 20 video-clips and served as the reference group for comparison of the readings of the group of MDR. Comparison between percentages of correct answers of MDR and the controls was done with a test (program WINPEP, software version 7.65). Results Mean diagnostic accuracy of readings of the MDR readers group A and 40 quality controlled cardiologists group B was, respectively 78% ± 0.12 and 99% ± 0.05 as compared to gold standard. Differences between mean values group A - group B was -17.00% (p=0.005; 95% IC = -25.93 - 13.07).

Another comparison, now made only between group of residents vs gold standard readings, showed that the most disagreement was evidenced by higher dispersion among readers in zones with more accentuated fluid accumulation, and was smaller in regions with normal patterns of A-lines as demonstrated in balloon graph (see figure). For the images with 0 B-lines most of the answers of MDR were in concordance with the gold standard and with more B-Lines the concordance was smaller. Conclusions B-line assessment is simple but a standardized training, specific certification and audit are warranted to harmonize reading criteria and optimize the potentially important clinical impact of the technique.