Bedside Evaluation of Critical Dyspnea. Proposal of a Diagnostic Score for Acute Heart Failure in Emergency Setting

Abstract

Aim: Dyspnea is a common cause of hospital admittance. Preliminary investigations in the pre-hospital phase and in the Emergency Department (ED) should detect the underlying cause. Time is crucial and emergency physicians have few diagnostic tools to manage patients presenting shortness of breath. In this study we assess the performance of lung ultrasound as a diagnostic tool in the evaluation of acute heart failure (AHF), in order to formulate an ultrasound-based diagnostic score.

Methods and Results: Over a two-year period, 236 consecutive patients admitted to our ED for non-traumatic dyspnea were enrolled in the study. All patients underwent lung ultrasound (LU) systematic evaluation reporting B-lines, in addition to standardized work-up. ROC curve showed an AUC=82.3% (95% CI=76.3%-87.9%) and AUC=75.5 (95% CI=68.4%-81.3%) for LU and NT pro-BNP respectively. About 18% of the patients enrolled showed a number of B-Lines >18, which were used as an early diagnostic test to detect patients with AHF: A sensitivity and specificity of 39.8% and 97.0% were found respectively. A score model was designed to diagnose the remaining patients including LU, chest X-Ray and NT-proBNP to supply high diagnostic accuracy (AUC=91.7%).

Conclusion: As known, LUS can be a useful tool for a prompt and accurate detection of AHF, allowing chest X-ray and biomarker evaluation to be avoided in a remarkable portion of dyspneic patients, which include about 40% of the actual AHF occurrences. In this way it is possible to reach an accurate diagnosis in a short amount of time, making it possible to start therapy precociously. An integrated approach that includes chest X-ray and NT-proBNP can improve diagnostic capabilities. Our proposed operative protocol minimizes the ratio between time of medical intervention and diagnostic accuracy of AHF, in patients presenting shortness of breath, starting from prehospital phase.

Keywords: Diagnosis; Heart failure; Ultrasound; Dyspnea; Emergency setting; Chest x-ray; Natriuretic peptides

Introduction

Dyspnea is one of the most common causes of admission to the emergency department worldwide: Acute heart failure (AHF) is a major cause of serious morbidity and death in the general population and one of the most common medical causes of hospitalization among people aged over 60 [1]. The incidence rate is significantly higher in men than in women and increases with age from 1.4/1000 person per year in subjects aged 55-59 to 47.4/1000 person per year in those aged 90 or over in Europe [2]. The age-adjusted average prevalence of AHF in the United States is 36 cases per 100,000 population and accounts for 10,000 deaths annually [3].
In clinical practice this symptomatology is usually investigated in pre-hospital phase only through anamnensis and physical examination; in the Emergency Department (ED) blood gas analysis (BGA), laboratory tests and chest X-ray can be performed as primary exams. BNP and NT pro-BNP are considered reliable biochemical markers to distinguish cardiogenic etiology from pulmonary, both for their diagnostic and prognostic value [4-6]. On the other hand, these biomarkers are affected by a "grey-zone" of uncertainty [7], they are not available in all hospitals and their dosage samples are expensive: thus our study aims to propose other tools to support the diagnostic process [8,9]. It is widely reported in literature that the ultrasound method can identify the presence of interstitial and pulmonary alveolar syndrome by noting the comet tail aspect of the artefacts, called B-lines, with high sensitivity and specificity [10-13]. This is an easy technique to perform and interpret, readily available and repeatable over time. Moreover it can be carried out at the bedside even in emergency situations [14,15]. The execution of this investigative technique in the ED and also in pre-hospital setting would provide additional and reliable information to patients with dyspnea favoring rapid identification of the etiology [16]. The main aim of this study is to identify a protocol model or management of acute dyspnea on the basis of lung ultrasound pattern, potentially combined with other validated diagnostic tools, not only in ED but also in pre-hospital phase.

Methods

Setting and participants

In a previous study, a comparison was made between lung ultrasound, chest X-ray and NT pro-BNP performance [13].

In this current monocentric, prospective, randomized study, we reassessed the data of the aforementioned study. We had enrolled all patients (Age>18 years) admitted to our Emergency Department over a two year period, with acute non-traumatic dyspnea. For organizational purposes we enrolled all consecutive patients every day for a period of 2 h a day. We included both patients brought in by ambulance and patients arriving autonomously. Exclusion criteria were: age<18 years, dyspnea following trauma, history of chronic interstitial lung disease (because of confounding pattern at LUS) [17].

The Local Ethical Committee of our Institute approved the study protocol, according to the Helsinki Declaration and all participants were able to provide informed consent for both parts of the study.

Diagnostic procedures

After triage, standardized diagnostic work-up included: 1) brief patient anamnensis (age, gender, symptoms, medical history); 2) Vital Parameters (Blood Pressure, Heart Rate, Arterial Oxygen Saturation, Respiratory Rate, Body Temperature); 3) 12 lead ECG; 4) Standard laboratory assessment (NT-proBNP-ECLIA, Roche Methodics*, Creatinine, BUN, C-Reactive Protein, Full Blood Count, Electrolytes); 5) Blood Gas Analysis. Data were collected and stored in a database.

Lung Ultrasound was performed immediately after by an independent operator (unaware of patient history and vitals), then the attending physician acquired full medical history, Physical Examination (according to Boston Criteria for Heart Failure) and requested Chest X-Ray (carried out by an expert Radiologist). Echocardiography was not considered in this study because not available 24 h/24 h. All LUS operators were EPs with specialized training in US or departmental researchers, EM residents who had undergone didactic training (1 h) and hands-on instruction (2 h) on thoracic US. Each resident then demonstrated proficiency in performing and interpreting a minimum of ten scans under direct attending supervision [18].

Lung ultrasound

We used Esaote MyLab30™ and MyLab70™ Ultrasound with Variable-Band Convex Array (3.5-5 MHz). Six transversal scans for each hemithorax (second and fourth intercostal space on the hemiclavear line, anterior axillar line, middle axillar line, see Figure 1). Basal scans of the lung were sampled in order to point out pleural effusions.

B-lines are defined as vertical, hyperechogenic, reverberation artefacts which arise from the pleural line to the bottom of the screen and move synchronously with lung sliding [19-21], a positive region is defined by the presence of three or more B-lines in a longitudinal plane between two ribs [22,23] Acute interstitial pulmonary syndrome was defined, according to the International Consensus Conference on LUS [24] as the presence of two or more positive regions for each hemithorax.

Diagnostic outcome (Reference standard)

We chose to use a consensus diagnosis as formal reference standard, in analogy with earlier studies and as recommended by recent diagnostic research guidelines [25,26]. An independent operator (unaware of patient history and vitals), then the attending physician acquired full medical history, Physical Examination (according to Boston Criteria for Heart Failure) and requested Chest X-Ray (carried out by an expert Radiologist). Echocardiography was not considered in this study because not available 24 h/24 h. All LUS operators were EPs with specialized training in US or departmental researchers, EM residents who had undergone didactic training (1 h) and hands-on instruction (2 h) on thoracic US. Each resident then demonstrated proficiency in performing and interpreting a minimum of ten scans under direct attending supervision [18].
expert panel judged all the available diagnostic information from each patient to determine the final diagnosis.

The outcome panel consisted of 1 cardiologist and 1 emergency physician. The approach used to reach a final diagnosis was by discussion of the whole material regarding each patient and then by consensus. The panel assessed each patient for the presence of heart failure following the criteria and approach outlined by ESC 2016 Heart Failure Guidelines [26].

Statistical analysis and score model design

Multivariate stepwise logistic regression analysis was carried out to identify, among all risk factors (independent variables), a statistically significant minimum subset of factors with the highest possible accuracy to predict acute heart failure. Logistic discrimination is generally preferable to linear discrimination in small samples, especially when data distribution is suspected to be non-Gaussian [27]. In the classical logistic regression, the dependent variable (AHF) assumes binary values, while both qualitative and quantitative independent variables (predictors) can be included in the model. In the stepwise process, one independent variable was added to, or removed from, the discriminant model at each step, on the basis of the criterion of maximum likelihood ratio. The process ends when no other statistical significant variables can be entered or removed.

Diagnostic accuracy was evaluated by sensitivity and specificity values for dichotomous tests and by the area (AUC) under ROC curve for quantitative tests. The Hosmer-Lemeshow goodness-of-fit test was used to evaluate model calibration. The so-designed multivariate logistic model was then transformed into a score model by rounding the ratio between the regression coefficients and the reference coefficient of lung ultrasound B-Lines to the nearest integer, so that the latter become equal to 1 (one point for each line B). This allowed avoiding a computer for the immediate diagnosis of AHF by a simple calculation of the total risk score and its comparison with a suitable cut-off value [28].

The Student t-test or the Mann-Whitney tests were used to compare quantitative Gaussian or non-Gaussian variables, respectively, to identify significant differences between patients with and without AHF. The Anderson-Darling statistical test was used to verify normality of data distribution. A statistical comparison between diagnostic tests was performed by evaluating the 95% confidence interval (CI) of sample estimates, by using a bootstrap approach [29].

The SPSS software (version 10) and Matlab package were used for statistical computations and score model design, respectively.

Results

Over 18-months, a total of 255 caucasian patients were evaluated, 19 of them were excluded because of a history of chronic interstitial lung disease. Of 236 patients enrolled, 117 were male (49.6%) and 119 female (50.4%) with an average age of 79.98 (SD: ± 12.13). 74 had a medical history of cardiac diseases (chronic heart failure, valvular diseases, ischemic heart disease), 38 had pulmonary disease (mainly COPD) and 63 both. In 65 patients (27.5% of the total) empiric therapy was administrated in pre-hospital setting. The final diagnoses are reported in Table 1 and continuous variables’ description is reported in Table 2.

All patients were evaluated with lung ultrasound (LUS) on admission with an average execution time of 3.75 ± 1.8 min. On the other hand time-to-chest X-ray was 59 ± 39 min (including the time needed to complete the report).

In this study we used the definition of the International Consensus Conference on Lung Ultrasound to determine interstitial pulmonary syndrome patterns [24]. We evaluated six scans for each hemithorax reporting the number of B-Lines. Figures 2A and 2B show the ROC curve of LUS estimated by evaluating the number of the B-Lines for the six zones on the right side and on the left. A high level of accuracy was exhibited in detecting AHF (AUC=82.3%; 95% CI=76.3%-87.9%); using only four scans (zone 3 to 6), instead of six, the estimated accuracy (AUC=83.3%) was not statistically different from the use of all six zones, because of a large 95% CI (77.5%-88.6%) of estimated AUC. LUS examination with only four scans seemed more convenient [13,24]. However, we preferred to consider six scans because they increased the rate of AHF patients correctly recognized only through LUS examination: further statistical analysis was made by using only six bilateral scans. In Figure 2A, ROC curve rises almost vertically.
from the left bottom corner, in correspondence to a very high specificity (above 97%), and begins to deviate to the right when the sensitivity is approximately equal to 40%. Therefore, we found that patients showing a number of LUS B-lines greater than 18 overall can be conveniently classified as AHF, without further investigation. We chose to use this number as cut-off value for the diagnosis of AHF. In accordance with this choice, we found: SE=39.8% (95% CI=28.6%-48.7%); SP=97.0% (95% CI=93.9%-99.3%); PPV=90.0% (95% CI=76.7%-96.7%); NPV=68.8% (95% CI=65.7%-70.3%). In particular, 18% of the whole group of patients was positive for LUS B-Lines, a percentage greater than the 15% that would have been obtained with only four scans. Among positive patients, only 10% (false positives) did not develop AHF. Among patients with B-lines ≤18, 87.5% of the whole sample, a percentage over 30% of false negative was observed. Therefore, negative patients need to be examined more closely.

In accordance with ESC HF 2016 guidelines [26], using a cut-off value of 300 pg/ml for NT pro-BNP to describe acute onset dyspnea, we found an overall sensitivity (SE) and specificity (SP) of 98.7% (95% CI=93.0%-99.9%) and 27.9% (95% CI=24.2%-28.6%), respectively. Negative and positive predictive values (NPV and PPV) resulted equal to 97.1% (95% CI=84.3%-99.9%) and 46.7% (95% CI=43.9%-47.2%). The ROC analysis showed a diagnostic accuracy for NT pro-BNP (AUC=75.5%; 95% CI=68.4%-81.3%) lower than that of LUS B-Lines.

The multivariate logistic regression model selected the most discriminant variables to be LUS B-Lines, chest X-Ray and NT pro-BNP. In order to derive an integer score from the logistic model, NT pro-BNP was conveniently discretized by grouping its values into the following three classes: 1) NT pro-BNP ≤300; 300<NT pro-BNP ≤2000; NT pro-BNP>2000. The score model was able to maximize diagnostic accuracy. Table 3 lists the partial scores, the sum of which forms the total individual model score. It allowed us to obtain an AUC=91.7% (95% CI=87.4%-94.7%), which can be considered highly accurate for first level exams in the acute phase. The ROC curve analysis of Figure 2B shows that the lower bound of 95% CI, which represents the worst statistical condition, is still highly accurate (AUC=87.4%). In order to minimize false negatives, we selected a high sensitivity point on the ROC curve, which corresponds to a cut-off score equal to 18, so that, for model score>18, we had: SE=93.4% (95% CI=87.0%-97.2%); SP=76.8% (95% CI=71.4%-78.7%); PPV=73.9% (95% CI=68.9%-76.9%); NPV=94.1% (95% CI=88.3%-97.5%) (Figure 3).

Table 3 Score model, named 18-LUNG. For each patient, the total score is obtained by adding the partial scores to be evaluated by the listed conditions.

Table 3

<table>
<thead>
<tr>
<th>Diagnostic 18-LUNG</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>500 pg/ml&lt;NT pro-BNP ≤2000 pg/ml</td>
<td>5</td>
</tr>
<tr>
<td>NT pro-BNP&gt;2000 pg/ml</td>
<td>12</td>
</tr>
<tr>
<td>X-Ray Alveolar Edema</td>
<td>25</td>
</tr>
<tr>
<td>X-Ray Interstitial Edema</td>
<td>14</td>
</tr>
<tr>
<td>X-Ray Pleural Effusion</td>
<td>10</td>
</tr>
<tr>
<td>LUS B-lines</td>
<td>1 pt. per line</td>
</tr>
</tbody>
</table>

AHF diagnosis is made if total score is >18 with 93.5% SE (95% CI=87.0%-97.8%) and 74.3% SP (95% CI=66.7%-82.5%)

Discussion

The statistical significance of this sample study proves that the simple and useful tool of LUS can be conveniently employed for the evaluation of dyspnea in the ED. Time-to-execution (<3 min) of LUS was shorter compared with chest-X ray (59 ± 39 min) [30,31]. Obviously patients with acute dyspnea need to be treated before the chest-X ray report, on the basis of clinical evaluation. On the other hand, with the help of ultrasound, it is possible to set the therapy more appropriately.

The accuracy using the minimum numbers of B-Lines required for the diagnosis of interstitial pulmonary syndrome (at least 12) [24] was basically good (AUC=84%). Four scans (3-6 zones bilaterally) in each hemithorax, instead of six scans, as previously proposed by many other studies, showed the same accuracy [32,33]. In case of pulmonary overload, in a supine patient, fluids tends to be distributed in lateral and posterior zones [34,35] thus allowing detection of interstitial syndrome on medium and posterior axillary regions rather than on paraclavicular region. In our study, the sensitivity of single LUS examination, with 12 B-Lines as diagnostic cut-off, shows a lower value than those previously reported in other studies [14], however this could be explained considering that many patients with AHF received diuretic therapy in pre-hospital setting. We selected a cut-off value higher than 12 B-Lines to reach the highest possible specificity

This article is available in: http://criticalcare.imedpub.com/archive.php
compatible with a not too marked reduction of sensitivity. Taking 18 B-lines as decision threshold, a high positive predictive value (90%) was obtained. This allowed us to minimize the number of patients (only four cases in our sample) who are unnecessarily treated for AHF, therefore allowing us to perform an immediate and early diagnosis of AHF using only LU for a large number of risk patients.

Our findings confirm the usefulness of NT pro-BNP in the role-out of AHF (NPV: 97.1%), but the cut-off value of 300 pg/ml [26], supplied too low specificity and positive predictive values (SP=27.9%; PPV=46.7%). Moreover, its diagnostic accuracy (AUC=75.5%) was considerably lower than LUS examination [36,37]. Higher values of NT pro-BNP have been associated with a wide variety of cardiac and non-cardiac causes with low diagnostic accuracy [38].

ROC curve analysis (Figures 2A and 2B) suggested a quantitative clinical protocol (Figure 3), named PLEURA18 (Protocol for Lung Evaluation by Ultrasound in Respiratory distress patients with Acute heart failure), which was able to reach a satisfactory compromise between diagnostic time reduction and accuracy. Considering 18 LUS B-Lines as a cut off, estimated sensitivity and specificity were respectively 39.8% and 97.0%. A sub-group of 40 patients (18% of the whole sample) had a number of LUS B-Lines greater than 18 and 36 of them developed AHF. Therefore, about 40% (95% CI=30%-50%) of the whole AHF patients could have been correctly identified by performing only the diagnostic test LUS B-Lines >18. Only 1.8% (95% CI=0.4%-3.8%) of patients (false positives) would be immediately and unnecessarily treated for AHF. Then, our protocol considers that patients exhibiting a number of LUS B-Lines>18 need not to perform Chest X-Ray in acute phase and were quickly diagnosed as AHF. Patients with a number of LUS B-Lines up to 18 (in our sample about 82%) must be evaluated by our proposed score model (Table 3) which involves Chest X-Ray and NT pro-BNP evaluation. The score model supplies high accuracy (AUC=91.7%) and, by choosing a decision threshold again equal to 18, it minimizes the number of false negative patients who would be mistakenly be considered not at risk of AHF (SE=93.4%, 95% CI=87.0%-97.2%; NPV=94.1%, 95% CI=88.3%-97.5%). The score model, so called 18-LUNG (Lung Ultrasound-NT pro-BNP-radiology), which combines two instrumental exams and one laboratory biomarker, has a superior diagnostic accuracy compared to each one of these taken alone [39,40].

Our clinical protocol was then completely determined by combining the early detection of several AHF patients by LUS B-Lines >18 and the diagnostic performance of score model on the remaining patients (having LUS B-Lines ≤ 18) more closely and extensively examined through Chest X-Ray and NT pro-BNP.

With respect to our sample data, the application of the proposed protocol gave 3% of false negative and 13.5% of false positive cases and correctly classified 83.5% of patients. A conspicuous part of patients (18% of the whole and 40% of those who actually developed AHF) were accurately diagnosed directly through LUS.

Finally, it should be reported that by choosing a lower cut-off value of the LUS B-Lines for the six-scans this diagnostic method can be used alone with an accuracy comparable to that of the Chest X-Ray (similar AUC values). Therefore, the LUS B-Lines technique is proposed as a significant aid to diagnosis of AHF, especially when combined appropriately with other tests such as the NT pro-BNP and Chest X-Ray, as in our proposed score model.

In conclusion, we propose a score to quantify the probability of AHF in dyspneic patients from ED, where LUS has an important role for its feasibility, reliability and cost-effective characteristics [41].

**Limitations**

The limitations of this study are primarily due to the absence of a testing group and the limited numbers of enrolled patients that causes large confidence intervals of statistical estimates. Moreover, in AHF management, there is no diagnostic “gold standard” to confront in emergency setting.

This should be considered a preliminary study where a clinical protocol management has been identified on the basis of collected data; further studies are needed to validate the protocol and confirm our findings.

**Conclusion**

LUS is a challenging instrument in emergency setting where time is crucial and rapid diagnosis is essential for the patients. Its accuracy to detect lung overload due to AHF can be a useful tool for the early identification of such patients admitted to ED for dyspnea. Moreover, the possibility to diagnose AHF using ultrasound alone, allows an early and targeted treatment even in pre-hospital phase.

In our study this diagnostic instrument played an important role in the management of dyspneic patients, so we suggest its utilization by Emergency Physicians: Lung Ultrasound is a rapid, easy and point of care tool.

Many studies compared LUS to other diagnostic tools [42,43] but our scoring model suggests an integrated and operative approach to patients with acute onset shortness of breath.

Furthermore, adopting our protocol, we could reduce time-to-treat and medical costs due to standard diagnostic management in almost one third of patients in our sample. In the remaining two thirds, an integrated approach with chest X-ray and NT pro-BNP can improve diagnostic capabilities in AHF.

In our ED we have started the validation of this protocol. We trust that other centers would like to verify our proposal including LUS in daily management of dyspneic patients.
References


