Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Sequential integration of pre-discharge lung ultrasound and the MEESSI-AHF score for safe emergency department discharge in decompensated heart failure: a prospective multicentre cohort study

Marta Torres Arrese

mtorresa@salud.madrid.org

Hospital Universitario Infanta Cristina

Pablo Rodríguez Fuertes

Hospital Universitario La Paz

Pablo Barberá Rausell

Hospital Universitario y Politécnico La Fe

Rocío Salas Dueñas

Hospital Universitario Fundación Alcorcón

Elena Espílez Laredo

Hospital Universitario Fundación Alcorcón

Gonzalo García de Casasola Sánchez

Hospital Universitario Fundación Alcorcón

Davide Luordo

Hospital Universitario Infanta Cristina

Geovanna Carolina Masapanta Caiza

Hospital Universitario Infanta Cristina

Yale Tung Chen

Hospital Universitario La Paz

Òscar Miró

Hospital Clínic Barcelona, IDIBAPS, Universitat de Barcelona

Pere Llorens

Hospital General Universitario Dr. Balmis, Universidad de Alicante

---

Research Article

Keywords: Heart failure, Emergency service, hospital, Ultrasonography, Patient readmission, Prognosis

Posted Date: November 18th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11537297/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background

Early revisit or readmission is common after direct emergency department (ED) discharge for decompensated heart failure (DHF, previously termed acute heart failure). We evaluated whether pre-discharge lung ultrasound, quantified as the number of lung fields with B-lines (NLFB-L), improves 30-day risk stratification, alone or integrated with the Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure (MEESSI-AHF).

Methods

Prospective multicentre cohort study of 96 adults with DHF discharged directly from the EDs of four Spanish tertiary hospitals between November 2023 and January 2026. The primary outcome was heart failure-related ED revisit or hospital readmission within 30 days. Associations were assessed with logistic regression (two-variable models), discrimination with the area under the receiver operating characteristic curve (AUC) compared by the DeLong test, and the optimal NLFB-L threshold with the Youden index. A sequential strategy stratified risk by pre-discharge congestion (NLFB-L < 5 vs ≥ 5), refined by baseline MEESSI-AHF category (low vs non-low).

Results

The primary outcome occurred in 26 patients (27.1%). NLFB-L was independently associated with the outcome after adjustment for natriuretic peptide levels (adjusted odds ratio 1.51 per additional lung field; 95% confidence interval (CI) 1.21–1.90; p < 0.001). NLFB-L discriminated better than the MEESSI-AHF score (AUC 0.849 vs 0.670; DeLong p = 0.0037), whereas combining both did not improve discrimination (AUC 0.853; p = 0.953). The optimal threshold was NLFB-L ≥ 5 (69.2% sensitivity, 82.9% specificity, 87.9% negative predictive value). NLFB-L < 5 carried a 12.1% event rate; among patients with NLFB-L ≥ 5, event rates were 35.7% with low-risk and 81.3% with non-low-risk MEESSI-AHF (odds ratio 31.42 versus NLFB-L < 5, 95% CI 7.32–134.85; p < 0.001).

Conclusions

Pre-discharge NLFB-L strongly predicted 30-day heart failure-related revisit or readmission. The added value of the MEESSI-AHF score emerged with sequential integration: a non-low MEESSI-AHF category refined prognosis only in patients with persistent congestion; when both are present, direct ED discharge should be reconsidered.

Trial registration:

Not applicable (observational study, not registered).

Background

Decompensated heart failure (DHF), the term that the 2026 European Society of Cardiology guidelines now recommend instead of acute heart failure [1], is a leading cause of emergency department (ED) visits, with substantial morbidity, mortality and resource use [2]. After stabilisation, roughly one-quarter to one-third of patients are discharged directly from the ED, entering a vulnerable phase with a high risk of heart failure (HF)-related revisit and readmission [2, 3].

Accurate risk stratification is therefore essential to ensure not only short-term survival but also sustained stability. The Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure (MEESSI-AHF) was validated to predict 30-day mortality [4] and is widely used in the ED [5, 6]. Clinical congestion scores such as the Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan (EVEREST) score were developed in hospitalised patients and were mainly linked to mortality or composite outcomes [7]. Yet short-term mortality and early post-discharge events are likely driven by different mechanisms: mortality largely reflects baseline vulnerability, comorbidity and end-organ dysfunction, whereas early revisit or readmission relates more to residual pulmonary congestion, which, assessed as B-lines by lung ultrasound (LUS), predicts adverse post-discharge outcomes in hospitalised patients [8–10].

Point-of-care LUS provides a rapid, non-invasive bedside assessment of residual pulmonary congestion [11]. B-lines, generated by increased extravascular lung water, detect pulmonary congestion more accurately than physical examination or chest radiography [12, 13]; however, evidence on their prognostic value after discharge comes almost exclusively from hospitalised patients [14, 15], and data on patients discharged directly from the ED remain scarce, despite this being the setting where such data may matter most. Whether, and how, pre-discharge LUS and an established ED risk score should be combined is therefore unknown: to our knowledge, no study has evaluated whether their integration improves risk stratification in patients discharged directly from the ED, a question directly relevant to safe ED discharge.

We aimed to evaluate whether pre-discharge LUS, quantified as the number of lung fields with B-lines (NLFB-L), improves 30-day risk stratification in patients with DHF discharged directly from the ED, alone or integrated with the MEESSI-AHF score. Rather than treating them as competing tools, we evaluated both within a sequential echo-clinical strategy in which pre-discharge LUS first identifies residual congestion and the MEESSI-AHF score, calculated at presentation, then contextualises baseline vulnerability.

Methods

Study design and setting

We conducted a prospective multicentre observational cohort study within the ICALTA project (from the Spanish Insuficiencia Cardiaca–ALTA, heart failure at ED discharge) in the EDs of four Spanish tertiary university hospitals: Hospital Universitario Fundación Alcorcón (Alcorcón, Madrid), Hospital Universitario Infanta Cristina (Parla, Madrid), Hospital Universitario La Paz (Madrid) and Hospital Universitario y Politécnico La Fe (Valencia). Patients were enrolled from 8 November 2023 to 21 January 2026. Reporting follows the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for cohort studies (Additional file 1). The study was approved by the Ethics Committee for Research with Medicines (CEIm) of Hospital Universitario Fundación Alcorcón (approval code 23/134; 2 November 2023) and by all participating centres, and was conducted in accordance with the Declaration of Helsinki. All participants provided written informed consent before inclusion. Data were pseudonymised and held in a secure institutional environment.

Participants

Adults (≥ 18 years) with DHF who were discharged directly home from the ED after initial assessment and treatment, without hospital admission, were eligible. DHF was diagnosed by the treating emergency physician, who was not a member of the study team, on the basis of new-onset or worsening symptoms and signs of HF; the diagnosis was retained for inclusion only when the pre-discharge cardiopulmonary ultrasound examination, performed by a study investigator, did not show an alternative diagnosis. Patients were excluded if they declined to participate or if they were haemodynamically unstable; patients in whom the ultrasound findings led to hospital admission were, by definition, not part of the cohort. Patients were enrolled consecutively during the periods in which a study investigator was available at each centre to perform the pre-discharge ultrasound examination; the numbers of eligible patients attending outside these periods, declining participation or excluded were not recorded.

Data collection and variables

Investigators prospectively recorded baseline demographic characteristics (age and sex), comorbidities (hypertension, diabetes mellitus, chronic lung disease, previous HF, HF admission during the previous year and follow-up in a specialised HF unit), baseline functional status assessed with the Barthel Index, vital signs at presentation and at discharge (mean arterial pressure and the peripheral oxygen saturation/fraction of inspired oxygen ratio, SpO_2/FiO_2), the intravenous furosemide dose administered in the ED, chronic and discharge HF treatments, and the EVEREST congestion score at discharge, using a standardised case report form. Laboratory variables at discharge included high-sensitivity troponin, N-terminal pro-B-type natriuretic peptide (NT-proBNP), carbohydrate antigen 125 (CA-125), creatinine, sodium and potassium, measured in the laboratories of the participating hospitals.

Pre-discharge cardiopulmonary ultrasound protocol

Pre-discharge cardiopulmonary ultrasound protocol

All patients underwent a bedside cardiopulmonary ultrasound examination immediately before ED discharge. The examination was performed at each centre by a study investigator, an emergency physician with formal training and more than six years' experience in point-of-care ultrasound who was not involved in the clinical care of the patient, using a Mindray M9 system (Shenzhen Mindray Bio-Medical Electronics, Shenzhen, China). LUS followed a 12-zone protocol (six lung fields per hemithorax: two anterior, two in the mid-axillary line and two posterior) [11]. The NLFB-L was recorded, a field being considered positive when three or more B-lines were present. Pleural effusion was graded on each side by its craniocaudal extent as 0 (absent), 1 (mild, one intercostal space), 2 (moderate, two or three intercostal spaces) or 3 (severe, four or more intercostal spaces), and a pulmonary congestion score (PCS), a study-defined index, was calculated as the sum of the NLFB-L (range 0–12) and the pleural effusion grades of both hemithoraces (range 0–6), giving a possible range of 0–18. Focused echocardiography assessed chamber dimensions, left ventricular ejection fraction (LVEF, visually estimated; reduced LVEF was defined as LVEF < 50%), diastolic function (including mitral E-wave velocity), the tricuspid annular plane systolic excursion (TAPSE; reduced if ≤ 17 mm), the estimated pulmonary artery systolic pressure (PASP), the TAPSE/PASP ratio and the echocardiographic probability of pulmonary hypertension, graded as low, intermediate or high from the peak tricuspid regurgitation velocity and additional signs from the ventricles, pulmonary artery and inferior vena cava according to the criteria of the European Society of Cardiology and the European Respiratory Society [16]. Systemic venous congestion was assessed using inferior vena cava diameter and collapsibility and the venous excess ultrasound (VExUS) score [17].

The ultrasound findings were recorded for research purposes only and were not communicated to the treating physicians, who made the discharge decision according to usual clinical practice, with the exception of unexpected high-risk findings, such as severe valvular heart disease or a previously unknown reduced LVEF, which were reported for patient safety. Because the examination was performed before discharge, the study investigators were necessarily unaware of the subsequent outcome.

Baseline clinical risk score

The MEESSI-AHF score was calculated at ED presentation using the original 13-variable model [4] and the MEESSI-AHF risk calculator [18], which provides the risk quintile (1, lowest risk, to 5, highest risk) and the corresponding risk category. The risk quintile was analysed as an ordinal variable (hereafter, the MEESSI-AHF score), and the risk category was dichotomised into low risk (quintiles 1–2) versus non-low risk (quintiles 3–5: intermediate, high and very high risk combined).

Primary outcome and follow-up

The primary outcome was the composite of HF-related ED revisit or hospital readmission for persistent or worsening HF within 30 days after ED discharge. Follow-up was performed at 30 days by the site investigator through review of the electronic health records and telephone contact with the patient.

Outcome events were adjudicated by the same investigator who had performed the ultrasound examination, who was therefore not blinded to the ultrasound findings; events were, however, objectively documented in the health records. Vital status was also recorded. Follow-up was complete for all patients. No deaths occurred within 30 days, so death was not a competing risk.

Statistical analysis

No formal a priori sample size calculation was performed; the sample size was determined by the number of eligible patients enrolled during the study period. Continuous variables were summarised as mean (standard deviation) or median (interquartile range) according to their distribution (Shapiro–Wilk test), and categorical variables as frequencies and percentages. Comparisons between patients with and without the primary outcome used the Welch t-test or the Mann–Whitney U test for continuous variables and the Pearson χ^2 or Fisher exact test for categorical variables; these comparisons were exploratory and not adjusted for multiple testing.

Logistic regression assessed the independent prognostic value of clinical, laboratory and ultrasound markers, focusing on the predictors of primary interest: the two LUS measures (NLFB-L and PCS), discharge NT-proBNP as the reference biomarker and the MEESSI-AHF score as the reference clinical tool. Given only 26 outcome events, multivariable analyses were restricted to two-variable models including the NLFB-L plus one competing predictor (log-transformed NT-proBNP or the dichotomised MEESSI-AHF category), maintaining approximately 13 events per variable, above the recommended minimum [19]. The explanatory power of each model was summarised with the Nagelkerke R^2, and calibration was assessed with the Hosmer–Lemeshow goodness-of-fit test. Discrimination was assessed with the area under the receiver operating characteristic (ROC) curve (AUC) and its 95% confidence interval (CI); the optimal NLFB-L threshold was identified with the Youden index, with 95% CIs for sensitivity, specificity and predictive values calculated by the Wilson method, and correlated ROC curves were compared with the DeLong test. For the ROC comparison, a combined model including the NLFB-L and the MEESSI-AHF score (risk quintile) was fitted, and its predicted probability was used as the test variable.

Finally, a sequential echo-clinical stratification model was constructed, with pre-discharge residual congestion (NLFB-L < 5 vs ≥ 5) as the primary stratifier. Among patients with NLFB-L < 5, the 30-day event rate did not differ by baseline MEESSI-AHF category (low vs non-low; Fisher exact test), so these patients were collapsed into a single reference group; this grouping was decided after inspection of the event rates and is therefore exploratory. Patients with NLFB-L ≥ 5 were then subdivided by MEESSI-AHF category, yielding three mutually exclusive risk groups. Odds ratios (ORs) with 95% CIs were estimated by logistic regression, using NLFB-L < 5 as the reference; absolute event rates with numerators and denominators are reported alongside the ORs. Analyses used SPSS version 30 (IBM Corp., Armonk, NY, USA), and figures were drawn with Matplotlib (Python 3); a two-sided $p < 0.05$ was considered significant.

Missing data

Missing values were infrequent (≤ 13% per variable; 46 values across 12 continuous variables in 19 of 96 patients). Values assumed to be missing at random were imputed in SPSS with a prespecified k-nearest neighbours algorithm (k = 5): similarity was defined from a 55-variable feature set, and each value was replaced by the mean of the five nearest patients, constrained to the observed range; missing discharge values were first replaced by admission values when available. Structural absences (events that did not occur) and physiological absences (e.g. the absent A wave in atrial fibrillation) were not imputed, because they represent values that do not exist rather than missing data. A complete-case sensitivity analysis was fully consistent with the imputed analyses.

Use of artificial intelligence

During the preparation of this work, the authors used Claude (Anthropic) to assist with language editing of the manuscript and with the formatting of the tables. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the published article. Artificial intelligence was not used to generate data or to perform the statistical analyses, and no generative artificial intelligence imagery was used.

Results

Study population and outcomes

Ninety-six patients with DHF discharged directly from the four EDs were included. Their baseline characteristics are shown in Table 1. The 30-day composite outcome occurred in 26 patients (27.1%): 25 (26.0%) had an HF-related ED revisit and 17 (17.7%) were readmitted to hospital (16 patients had both events). No patient died within 30 days.

Table 1
Characteristics of patients with and without the 30-day primary outcome

Variable	Total (n = 96)	No event (n = 70)	Event (n = 26)	p-value
Demographic characteristics				
Age (years)[a]	81.5 [76.3–87.0]	82.5 [77.8–87.0]	79.5 [74.5–85.3]	0.137
Male sex	45 (46.9%)	30 (42.9%)	15 (57.7%)	0.196
Comorbidities and history				
Barthel Index[a]	100 [80–100]	100 [80–100]	100 [87.5–100]	0.769
Hypertension	79 (82.3%)	61 (87.1%)	18 (69.2%)	0.068
Diabetes mellitus	32 (33.3%)	24 (34.3%)	8 (30.8%)	0.745
Chronic lung disease	17 (17.7%)	11 (15.7%)	6 (23.1%)	0.387
Previous HF	60 (62.5%)	43 (61.4%)	17 (65.4%)	0.722
HF admission during the previous year	28 (29.2%)	16 (22.9%)	12 (46.2%)	0.026
HF-unit follow-up	20 (20.8%)	10 (14.3%)	10 (38.5%)	0.010
Clinical variables at discharge				
Mean arterial pressure (mmHg)	90.4 (14.2)	91.6 (12.6)	87.1 (17.9)	0.247
SpO_2/FiO_2 ratio[a]	452.4 [442.9–457.1]	452.4 [442.9–457.1]	447.6 [436.9–457.1]	0.787
Laboratory variables at discharge				
Creatinine (mg/dL)[a]	1.12 [0.87–1.46]	1.08 [0.84–1.41]	1.28 [0.90–1.69]	0.183
Sodium (mmol/L)	139.5 (3.5)	139.8 (2.9)	138.7 (4.7)	0.266
Potassium (mmol/L)	4.00 (0.54)	4.05 (0.53)	3.89 (0.57)	0.221
High-sensitivity troponin (ng/L)[a]	22 [14–35]	21 [13–30.5]	23 [14–46.5]	0.332
NT-proBNP (pg/mL)[a]	2802 [1740–5823]	2442.5 [1660.5–4443.3]	5570.5 [2266–11342.8]	0.012
CA-125 (U/mL)[a]	29.4 [15.9–83.0]	26.3 [15.7–71.1]	34.6 [16.2–117.3]	0.245
Prognostic scores				

Variable	Total (n = 96)	No event (n = 70)	Event (n = 26)	p-value
Demographic characteristics				
EVEREST score at discharge[a]	3 [2–5]	3 [2–4]	4 [2–5]	0.179
MEESSI-AHF score (risk quintile)[a]	2 [1–3]	1 [1–2]	3 [1–3]	0.006
Non-low MEESSI-AHF risk	29 (30.2%)	15 (21.4%)	14 (53.8%)	0.002
Lung ultrasound				
NLFB-L[a]	4 [2–5]	3 [2–4]	6 [4–9]	< 0.001
Pulmonary congestion score[a]	6 [4–8]	5 [3–7]	8.5 [6–13]	< 0.001
Cardiac ultrasound				
Mitral E-wave velocity (cm/s)[a]	105.5 [87.0–123.8]	107.5 [86.8–128.0]	100.5 [86.3–112.3]	0.358
Reduced LVEF	26 (27.1%)	21 (30.0%)	5 (19.2%)	0.291
Reduced TAPSE	44 (45.8%)	32 (45.7%)	12 (46.2%)	0.969
PASP (mmHg)	38.6 (16.6)	37.5 (16.6)	41.7 (16.4)	0.272
TAPSE/PASP ratio[a]	0.45 [0.30–0.68]	0.45 [0.33–0.69]	0.45 [0.26–0.63]	0.542
High probability of pulmonary hypertension	43 (44.8%)	29 (41.4%)	14 (53.8%)	0.277
Systemic venous congestion				
Inferior vena cava diameter (cm)[a]	2.20 [1.50–2.50]	2.20 [1.50–2.50]	2.20 [1.88–2.60]	0.222
VExUS score[a]	1 [0–1]	1 [0–1]	1 [0–3]	0.064
Hepatic vein systolic flow reversal	23 (24.0%)	14 (20.0%)	9 (34.6%)	0.136
Portal vein pulsatility > 50%	24 (25.0%)	14 (20.0%)	10 (38.5%)	0.063
Intrarenal monophasic venous flow pattern	14 (14.6%)	6 (8.6%)	8 (30.8%)	0.018

Variable	Total (n = 96)	No event (n = 70)	Event (n = 26)	p-value
Demographic characteristics				

Data are mean (standard deviation), median [interquartile range] or n (%). The primary outcome was a 30-day heart failure-related emergency department revisit or hospital readmission. Between-group p-values were obtained with the Welch t-test (normally distributed continuous variables), the Mann–Whitney U test (non-normally distributed continuous variables) and the Pearson χ^2 or Fisher exact test (categorical variables); comparisons were exploratory and not adjusted for multiple testing. The SpO_2/FiO_2 ratio was calculated with FiO_2 expressed as a fraction. The MEESSI-AHF score is the risk quintile provided by the MEESSI-AHF calculator (1, lowest risk, to 5, highest risk); the non-low risk category corresponds to quintiles 3–5. SI conversion: creatinine in mg/dL × 88.4 = µmol/L (overall median 1.12 mg/dL = 99 µmol/L); sodium and potassium in mmol/L (numerically identical to mEq/L); troponin in ng/L; NT-proBNP in pg/mL (numerically identical to ng/L). Adult reference ranges: creatinine 0.6–1.2 mg/dL (53–106 µmol/L); sodium 135–145 mmol/L; potassium 3.5–5.1 mmol/L; high-sensitivity troponin < 30 ng/L; NT-proBNP < 125 pg/mL; CA-125 < 35 U/mL; biomarker reference limits are assay- and age-dependent and correspond to those of the participating laboratories. Reduced LVEF, visually estimated LVEF < 50%; reduced TAPSE, TAPSE ≤ 17 mm; high probability of pulmonary hypertension according to the echocardiographic criteria of the European Society of Cardiology and the European Respiratory Society. [a]Median [interquartile range] because of non-normal distribution. CA-125, carbohydrate antigen 125; EVEREST, Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan; FiO_2, fraction of inspired oxygen; HF, heart failure; LVEF, left ventricular ejection fraction; MEESSI-AHF, Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure; NLFB-L, number of lung fields with B-lines; NT-proBNP, N-terminal pro-B-type natriuretic peptide; PASP, pulmonary artery systolic pressure; SpO_2, peripheral oxygen saturation; TAPSE, tricuspid annular plane systolic excursion; VExUS, venous excess ultrasound.

Additional files

Compared with patients without events, those with events more often had a prior-year HF admission (46.2% vs 22.9%; p = 0.026), were more commonly followed in a specialised HF unit (38.5% vs 14.3%; p = 0.010), presented with a higher MEESSI-AHF score (median risk quintile 3 vs 1; interquartile range 1–3 vs 1–2; p = 0.006), and were more often classified into a non-low MEESSI-AHF risk category (53.8% vs 21.4%; p = 0.002). NT-proBNP levels at discharge were also higher (p = 0.012; Table 1).

[Table 1 near here]

Cardiopulmonary ultrasound and multivariable predictors

On pre-discharge ultrasound, patients with the primary outcome had a higher NLFB-L (median 6 vs 3; interquartile range 4–9 vs 2–4; p < 0.001) and PCS (median 8.5 vs 5; interquartile range 6–13 vs 3–7; p < 0.001) than those without. The intrarenal monophasic venous flow pattern was also more frequent in the event group (30.8% vs 8.6%; p = 0.018). None of the other variables shown in Table 1 differed significantly between groups.

In univariate logistic regression, the NLFB-L (OR 1.53 per additional field; 95% CI 1.25–1.86; p < 0.001), the PCS (OR 1.36 per point; 95% CI 1.18–1.57; p < 0.001), log NT-proBNP (OR 2.16; 95% CI 1.24–3.75; p = 0.006) and a non-low MEESSI-AHF risk category (OR 4.28; 95% CI 1.64–11.17; p = 0.003) were

significantly associated with the 30-day primary outcome (Table 2). In two-variable models, the NLFB-L remained the sole independent predictor after adjustment for the non-low MEESSI-AHF risk category (adjusted OR 1.47; 95% CI 1.19–1.82; p < 0.001) or log NT-proBNP (adjusted OR 1.51; 95% CI 1.21–1.90; p < 0.001), while both competing predictors lost significance (non-low MEESSI-AHF risk category: adjusted OR 1.89, 95% CI 0.59–6.08; p = 0.284; log NT-proBNP: adjusted OR 1.06, 95% CI 0.54–2.06; p = 0.875) (Table 2). The echo-analytical model (NLFB-L plus log NT-proBNP) was well calibrated (Hosmer–Lemeshow p = 0.14), whereas the echo-clinical model with the dichotomised non-low MEESSI-AHF category showed evidence of miscalibration (Hosmer–Lemeshow p = 0.02). Adding the non-low MEESSI-AHF category to the NLFB-L modestly increased explanatory power (Nagelkerke R^2 0.352 vs 0.339).

Table 2
Logistic regression models for the prediction of the 30-day primary outcome

Model	Predictor	OR	95% CI	p-value	Nagelkerke R^2
Univariate models					
Analytical model	log NT-proBNP	2.16	1.24–3.75	0.006	0.118
Clinical model	Non-low MEESSI-AHF risk	4.28	1.64–11.17	0.003	0.130
Lung ultrasound model (NLFB-L)	NLFB-L	1.53	1.25–1.86	< 0.001	0.339
Lung ultrasound model (PCS)	PCS	1.36	1.18–1.57	< 0.001	0.304
Multivariable models					
Echo-clinical model	NLFB-L	1.47	1.19–1.82	< 0.001	0.352
	Non-low MEESSI-AHF risk	1.89	0.59–6.08	0.284	
Echo-analytical model	NLFB-L	1.51	1.21–1.90	< 0.001	0.340
	log NT-proBNP	1.06	0.54–2.06	0.875	

The primary outcome was a 30-day heart failure-related emergency department revisit or hospital readmission. Each multivariable model included two predictors (approximately 13 events per variable): the echo-clinical model included the NLFB-L and the dichotomised MEESSI-AHF category, and the echo-analytical model included the NLFB-L and log-transformed NT-proBNP. Nagelkerke R^2 measures the explanatory power of each fitted model. The MEESSI-AHF category was dichotomised as low versus non-low risk (intermediate, high and very high). Adjusted estimates for both predictors of each

multivariable model are shown; in each model the competing predictor lost significance after adjustment for the NLFB-L. CI, confidence interval; MEESSI-AHF, Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure; NLFB-L, number of lung fields with B-lines; NT-proBNP, N-terminal pro-B-type natriuretic peptide; OR, odds ratio; PCS, pulmonary congestion score.

Receiver operating characteristic analysis and sequential echo-clinical risk stratification

ROC analysis showed an AUC of 0.849 (95% CI 0.772–0.925; p < 0.001) for the NLFB-L, 0.798 (95% CI 0.702–0.893; p < 0.001) for the PCS and 0.670 (95% CI 0.543–0.797; p = 0.009) for the MEESSI-AHF score (Fig. 1 and Additional file 2: Table S1). The NLFB-L showed the highest discrimination among the individual predictors, significantly outperforming both the MEESSI-AHF score (DeLong p = 0.0037) and discharge NT-proBNP (AUC 0.672, 95% CI 0.540–0.805; DeLong p = 0.0003), whereas it did not differ significantly from the PCS (DeLong p = 0.778). Adding the MEESSI-AHF score (risk quintile) to the NLFB-L did not improve discrimination over the ultrasound-only model (combined AUC 0.853, 95% CI 0.776–0.930; p < 0.001; DeLong p = 0.953) (Fig. 1).

The optimal NLFB-L threshold was ≥ 5 lung fields, yielding 69.2% sensitivity (95% CI 50.0–83.5%), 82.9% specificity (95% CI 72.4–89.9%), 60.0% positive predictive value (95% CI 42.3–75.4%) and 87.9% negative predictive value (95% CI 77.9–93.7%). At this threshold, patients with NLFB-L ≥ 5 had a significantly higher 30-day outcome rate than those with < 5 (60.0%, 18/30, vs 12.1%, 8/66; OR 10.88, 95% CI 3.85–30.74; p < 0.001).

To evaluate the sequential strategy, patients were first classified by residual pulmonary congestion (NLFB-L < 5 vs ≥ 5). Because event rates among patients with NLFB-L < 5 were similar regardless of MEESSI-AHF category (13.2% vs 7.7%; Fisher exact p = 1.000), they were combined into a single reference group (12.1%, 8/66). Among patients with NLFB-L ≥ 5, baseline MEESSI-AHF category further stratified risk (Table 3 and Fig. 2). Within this group, the event rate was 35.7% (5/14) in patients with a low-risk MEESSI-AHF category (OR 4.03 versus NLFB-L < 5, 95% CI 1.08–15.07; p = 0.039) and reached 81.3% (13/16) in those with a non-low-risk category (OR 31.42, 95% CI 7.32–134.85; p < 0.001). The overall comparison across the three groups was significant (Pearson χ^2 = 31.78, 2 degrees of freedom; p < 0.001). This three-group model achieved the highest explanatory power among all evaluated models (Nagelkerke R^2 = 0.386).

Table 3
Sequential echo-clinical risk stratification for the 30-day primary outcome

Group	n	Events	OR	95% CI	p-value
NLFB-L < 5	66	8 (12.1%)	1 (reference)	—	—
NLFB-L ≥ 5 + low-risk MEESSI-AHF	14	5 (35.7%)	4.03	1.08–15.07	0.039
NLFB-L ≥ 5 + non-low-risk MEESSI-AHF	16	13 (81.3%)	31.42	7.32–134.85	< 0.001

The primary outcome was a 30-day heart failure-related emergency department revisit or hospital readmission. Because event rates among patients with NLFB-L < 5 were similar regardless of baseline MEESSI-AHF risk (13.2% vs 7.7%; Fisher exact p = 1.000), these subgroups were combined into a single reference category. The overall comparison across the three groups was significant (Pearson $\chi^2 = 31.78$, 2 degrees of freedom; $p < 0.001$). Odds ratios and 95% CIs were derived from a logistic regression model using NLFB-L < 5 as the reference. This model achieved the highest explanatory power among all evaluated models (Nagelkerke $R^2 = 0.386$). CI, confidence interval; MEESSI-AHF, Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure; NLFB-L, number of lung fields with B-lines; OR, odds ratio.

Discussion

In this prospective multicentre cohort of patients with DHF discharged directly from the ED, three findings emerged. First, residual pulmonary congestion at discharge, objectively quantified by the NLFB-L, was a strong independent predictor of 30-day HF-related ED revisit or hospital readmission. Second, the stand-alone MEESSI-AHF score showed limited discrimination for this early, non-fatal outcome, consistent with its design for short-term mortality rather than fluid-driven events. Third, integrating the two enabled sequential refinement: within the subgroup with significant residual congestion (NLFB-L ≥ 5), the MEESSI-AHF category further stratified risk.

Our findings extend to patients discharged directly from the ED evidence that was previously limited to hospitalised cohorts. Coiro et al. [9] and Gargani et al. [10] showed that persistent pre-discharge B-line congestion identifies a markedly higher short-term risk of readmission and death, and the meta-analysis by Suhardi et al. [15] found residual congestion associated with roughly three-fold higher mortality and four-fold higher rehospitalisation, strongest at short-term (< 6 months) follow-up, consistent with our 30-day window.

The novelty lies in applying this concept to patients discharged directly from the ED, a time-pressured setting requiring rapid, objective decisions [20]. Whereas decongestion has traditionally been judged by insensitive signs such as absent crackles [12], B-lines reflect extravascular lung water at a subclinical,

haemodynamically silent stage. Here, NLFB-L ≥ 5 was associated with a 30-day event rate of 60.0%, versus 12.1% in ‘dry’ patients, while preserving a high negative predictive value (87.9%).

The limited stand-alone discrimination of the MEESSI-AHF score for early, non-fatal events does not undermine this validated tool but reflects its original design: it was validated to predict 30-day mortality, with excellent discrimination [4]. In earlier ED cohorts, its predictive capacity fell for non-fatal events such as 30-day revisit or readmission after direct discharge, with reported ROC areas around 0.54 [5, 6], consistent with the modest discrimination observed in our series (AUC 0.670). This divergence is probably pathophysiological: short-term mortality reflects baseline reserve, comorbidity and end-organ damage, captured by the MEESSI-AHF score [4], whereas early revisit or readmission is driven mainly by residual, often subclinical, congestion, for which clinical scores are inherently limited, underscoring the complementary role of LUS [9, 10, 15].

The absence of improvement in overall discrimination (DeLong p = 0.953) should not be read as a lack of clinical utility: the NLFB-L alone captures most of the prognostic information on residual congestion, the principal correlate of early post-discharge events, and the MEESSI-AHF score adds value only once congestion is identified. In the ED, disposition requires simple, sequential algorithms rather than marginal gains in AUC; indeed, adding the MEESSI-AHF risk quintile to the NLFB-L did not improve discrimination, and the dichotomised category, when added to the NLFB-L, showed evidence of miscalibration, suggesting that its value lies in sequential application to patients with residual congestion rather than in additive discrimination across the whole cohort. Accordingly, the baseline MEESSI-AHF category did not detectably stratify risk among patients with NLFB-L < 5, whereas it separated two distinct phenotypes among those with NLFB-L ≥ 5.

Clinically, these findings support moving beyond physical examination towards objective bedside assessment of residual congestion before ED discharge [21]. The Second Universal Definition of Heart Failure likewise emphasises confirming congestion objectively, including by LUS [22]. Indeed, up to 40% of patients deemed clinically ‘dry’ have subclinical congestion at discharge, associated with worse outcomes [23]. A pragmatic approach would be to regard NLFB-L ≥ 5 as a ‘red flag’ warranting reassessment of the discharge decision (for example, further decongestion or a period of observation), applying the MEESSI-AHF category only thereafter; this strategy requires prospective evaluation. Patients with NLFB-L ≥ 5 and a non-low MEESSI-AHF category had an exceptionally high event rate (81.3%), so direct discharge home should be reconsidered in this subgroup.

Our cohort comprised older adults (median age 81.5 years), with a high prevalence of previous HF, and was recruited in tertiary hospitals with established point-of-care ultrasound programmes; the proposed strategy therefore applies most directly to similar ED populations and requires confirmation in other settings.

Limitations

First, because this was an observational study, residual confounding cannot be excluded and the associations do not establish causality; other variables associated with the outcome in univariate comparisons (previous HF admission, HF-unit follow-up and the intrarenal venous flow pattern) could not be entered into multivariable models given the number of events. Second, patients were enrolled consecutively only during the periods in which a trained investigator was available to perform the ultrasound examination, the number of eligible patients not enrolled was not recorded, and recruitment was unevenly distributed across centres, so selection bias cannot be excluded. Third, the small sample and limited number of events (n = 26) required parsimonious two-variable models; although the events-per-variable ratio remained adequate (> 10), confidence intervals were wide in smaller subgroups, particularly the three-group stratification, which was data-driven and should be considered exploratory until externally validated; calibration tests also have limited power in samples of this size. Fourth, the optimal NLFB-L threshold (≥ 5) was derived with the Youden index in the same cohort and also requires external validation. Fifth, inter- and intra-observer variability was not assessed [24], and scans were performed by experienced emergency physicians, potentially limiting generalisability to less experienced centres [25]. Sixth, outcome ascertainment was performed by the investigator who had performed the ultrasound examination and was therefore not blinded to its results, although the outcome events were objectively documented in the health records. Seventh, sex-specific differences were not examined, although these have been described in this population [26]. Finally, the outcome differed from that for which the MEESSI-AHF score was validated (30-day mortality); this is therefore not a revalidation but an assessment of its sequential role alongside discharge LUS.

Conclusions

In patients with DHF discharged directly from the ED, pre-discharge LUS should be the first step in risk stratification, as residual congestion was the strongest predictor of 30-day post-discharge events, and an NLFB-L < 5 may serve as a practical bedside criterion supporting discharge, pending external validation. These findings support a sequential echo-clinical strategy: LUS identifies residual congestion, after which the MEESSI-AHF score contextualises systemic vulnerability, isolating a very-high-risk subgroup for whom direct discharge home should be reconsidered.

Abbreviations

AUC
area under the receiver operating characteristic curve
CA-125
carbohydrate antigen 125
CEIm
Ethics Committee for Research with Medicines (Comité de Ética de la Investigación con medicamentos)
CI
confidence interval
DHF

decompensated heart failure
ED
emergency department
EVEREST
Efficacy of Vasopressin Antagonism in Heart Failure Outcome Study with Tolvaptan
FiO_2
fraction of inspired oxygen
HF
heart failure
ICALTA
Insuficiencia Cardiaca–ALTA (heart failure at emergency department discharge
project acronym)
LUS
lung ultrasound
LVEF
left ventricular ejection fraction
MEESSI-AHF
Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure
NLFB-L
number of lung fields with B-lines
NT-proBNP
N-terminal pro-B-type natriuretic peptide
OR
odds ratio
PASP
pulmonary artery systolic pressure
PCS
pulmonary congestion score
ROC
receiver operating characteristic
SpO_2
peripheral oxygen saturation
STROBE
Strengthening the Reporting of Observational Studies in Epidemiology
TAPSE
tricuspid annular plane systolic excursion
VExUS
venous excess ultrasound.

Declarations

Declarations

Ethics approval and consent to participate

The study was approved by the Ethics Committee for Research with Medicines (CEIm) of Hospital Universitario Fundación Alcorcón, Madrid, Spain (approval code 23/134; date of approval 2 November 2023) and by all participating centres, and was conducted in accordance with the Declaration of Helsinki. Written informed consent to participate was obtained from all participants before inclusion.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' information

MTA is a specialist in internal medicine and in emergency medicine and coordinates the Clinical Ultrasound Unit of the Emergency Department of Hospital Universitario Infanta Cristina, Parla, Madrid, Spain.

Tables

Additional file 1

Additional_file_1.docx (Microsoft Word document). **STROBE checklist.** Completed Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) checklist for cohort studies, indicating the page and line of the manuscript where each item is reported.

Additional file 3

Additional_file_3.xlsx (Microsoft Excel workbook). **De-identified patient-level dataset.** De-identified 30-day dataset of the 96 patients included in the study (one row per patient), in English, with a README sheet and a data dictionary describing every variable, its original case report form name, its units and its coding. Calendar dates, clinical record numbers and follow-up beyond 30 days have been removed.

Funding

The project was awarded a research grant by the Fundación Española de Medicina de Urgencias y Emergencias (FEMUE; Spanish Foundation of Emergency Medicine) to support the costs of the study, including publication costs; no grant number was assigned. No other specific funding was received. The funder had no role in the design of the study, in the collection, analysis and interpretation of the data, or in the writing of the manuscript.

Author Contribution

MTA conceived and designed the study, coordinated the project, acquired and curated the data, performed the statistical analysis and drafted the manuscript. PRF, PBR, RSD and EEL acquired and curated the data at their centres and critically revised the manuscript. DL prepared the figures and critically revised the manuscript. GCMC performed the literature review and critically revised the manuscript. GGCS, YTC and ÒM validated the results, contributed to the supervision of the study and critically revised the manuscript. PL validated the results, led the supervision of the study and critically revised the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors thank the Fundación Española de Medicina de Urgencias y Emergencias (FEMUE) for the research grant awarded to this project.

Data Availability

All data supporting the findings of this study are available within the article and its additional files. The de-identified patient-level dataset generated and analysed during the current study is provided as Additional file 3 (Microsoft Excel workbook with a README sheet and a data dictionary). Further information is available from the corresponding author on reasonable request.

References

1. Køber L, Adamo M, Ruwald AC, Tomasoni D, Anderson LJ, Andersson C, et al. 2026 ESC Guidelines for the management of heart failure. Eur Heart J. 2026;47:5346–457.
2. Miró Ò, Peacock FW, McMurray JJ, Bueno H, Christ M, Maisel AS, et al. European Society of Cardiology–Acute Cardiovascular Care Association position paper on safe discharge of acute heart failure patients from the emergency department. Eur Heart J Acute Cardiovasc Care. 2017;6:311–20.

3. Greene SJ, Fonarow GC, Vaduganathan M, Khan SS, Butler J, Gheorghiade M. The vulnerable phase after hospitalization for heart failure. Nat Rev Cardiol. 2015;12:220–9.
4. Miró Ò, Rossello X, Gil V, Martín-Sánchez FJ, Llorens P, Herrero-Puente P, et al. Predicting 30-day mortality for patients with acute heart failure in the emergency department: a cohort study. Ann Intern Med. 2017;167:698–705.
5. Rossello X, Bueno H, Gil V, Jacob J, Martín-Sánchez FJ, Llorens P, et al. MEESSI-AHF risk score performance to predict multiple post-index event and post-discharge short-term outcomes. Eur Heart J Acute Cardiovasc Care. 2021;10:142–52.
6. Miró Ò, Gil V, Rosselló X, Martín-Sánchez FJ, Llorens P, Jacob J, et al. Patients with acute heart failure discharged from the emergency department and classified as low risk by the MEESSI score (multiple risk estimate based on the Spanish emergency department scale): prevalence of adverse events and predictability. Emergencias. 2019;31:5–14.
7. Ambrosy AP, Pang PS, Khan S, Konstam MA, Fonarow GC, Traver B, et al. Clinical course and predictive value of congestion during hospitalization in patients admitted for worsening signs and symptoms of heart failure with reduced ejection fraction: findings from the EVEREST trial. Eur Heart J. 2013;34:835–43.
8. Dubón-Peralta EE, Lorenzo-Villalba N, García-Klepzig JL, Andrès E, Méndez-Bailon M. Prognostic value of B lines detected with lung ultrasound in acute heart failure. A systematic review. J Clin Ultrasound. 2022;50:273–83.
9. Coiro S, Rossignol P, Ambrosio G, Carluccio E, Alunni G, Murrone A, et al. Prognostic value of residual pulmonary congestion at discharge assessed by lung ultrasound imaging in heart failure. Eur J Heart Fail. 2015;17:1172–81.
10. Gargani L, Pang PS, Frassi F, Miglioranza MH, Dini FL, Landi P, et al. Persistent pulmonary congestion before discharge predicts rehospitalization in heart failure: a lung ultrasound study. Cardiovasc Ultrasound. 2015;13:40.
11. 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.
12. Drum B, La Course B, Kelly M, York A, Worrall E, Martins J, et al. Does this patient have volume overload? The rational clinical examination. JAMA. 2026;335:1159–68.
13. McGivery K, Atkinson P, Lewis D, Taylor L, Harris T, Gadd K, et al. Emergency department ultrasound for the detection of B-lines in the early diagnosis of acute decompensated heart failure: a systematic review and meta-analysis. CJEM. 2018;20:343–52.
14. Cohen A, Li T, Maybaum S, Fridman D, Gordon M, Shi D, et al. Pulmonary congestion on lung ultrasound predicts increased risk of 30-day readmission in heart failure patients. J Ultrasound Med. 2023;42:1809–18.
15. Suhardi KF, Prasetyo A, Malik R, Pauliana, Viazelda AT. Quantification of residual pulmonary congestion defined by B-line findings on lung ultrasound to predict cardiovascular events in acute

heart failure: a systematic review and meta analysis. Indian Heart J. 2025;77:258–66.

16. Humbert M, Kovacs G, Hoeper MM, Badagliacca R, Berger RMF, Brida M, et al. 2022 ESC/ERS Guidelines for the diagnosis and treatment of pulmonary hypertension. Eur Heart J. 2022;43:3618–731.
17. Beaubien-Souligny W, Rola P, Haycock K, Bouchard J, Lamarche Y, Spiegel R, et al. Quantifying systemic congestion with Point-Of-Care ultrasound: development of the venous excess ultrasound grading system. Ultrasound J. 2020;12:16.
18. Sociedad Española de Medicina de Urgencias y Emergencias (SEMES), ICA-SEMES Research Group. MEESSI-AHF risk model calculator. 2018. http://meessi-ahf.risk.score-calculator-ica-semes.portalsemes.org/. Accessed 26 Oct 2026.
19. Peduzzi P, Concato J, Kemper E, Holford TR, Feinstein AR. A simulation study of the number of events per variable in logistic regression analysis. J Clin Epidemiol. 1996;49:1373–9.
20. Miró Ò, Rodríguez-Cabrera M, Alquézar-Arbé A, Piñol C, González del Castillo J, Burillo-Putze G, et al. Emergency department length of stay of hospitalized patients and short-term outcomes (EDEN-18 study). Emergencias. 2025;37:343–52.
21. Bernaudo D, Guerrero-León MLÁ, Rodríguez-Cortés M, Tártalo-Hernández D, de Mora-Alfaro R, Landete Martín L, et al. Diagnostic algorithm for urgent causes of dyspnea based on lung auscultation and ultrasound: the CLINIC-LUS prospective study. Emergencias. 2026;38:123–30.
22. Walsh MN, Køber L, Sliwa K, Adamo M, Agarwal A, Banerjee A, et al. AHA/ACC/ESC/WHF expert consensus document: second universal definition of heart failure (2026). Circulation. 2026;154:e279–93.
23. Rivas-Lasarte M, Maestro A, Fernández-Martínez J, López-López L, Solé-González E, Vives-Borrás M, et al. Prevalence and prognostic impact of subclinical pulmonary congestion at discharge in patients with acute heart failure. ESC Heart Fail. 2020;7:2621–8.
24. Gullett J, Donnelly JP, Sinert R, Hosek B, Fuller D, Hill H, et al. Interobserver agreement in the evaluation of B-lines using bedside ultrasound. J Crit Care. 2015;30:1395–9.
25. Russell FM, Ferre R, Ehrman RR, Noble V, Gargani L, Collins SP, et al. What are the minimum requirements to establish proficiency in lung ultrasound training for quantifying B-lines? ESC Heart Fail. 2020;7:2941–7.
26. Fuentes J, Cabello I, Roset A, López-Delmas N, Malchair P, Jacob J. Outcomes in older adults with acute heart failure according to sex. Emergencias. 2026;38:234–5.

Figures

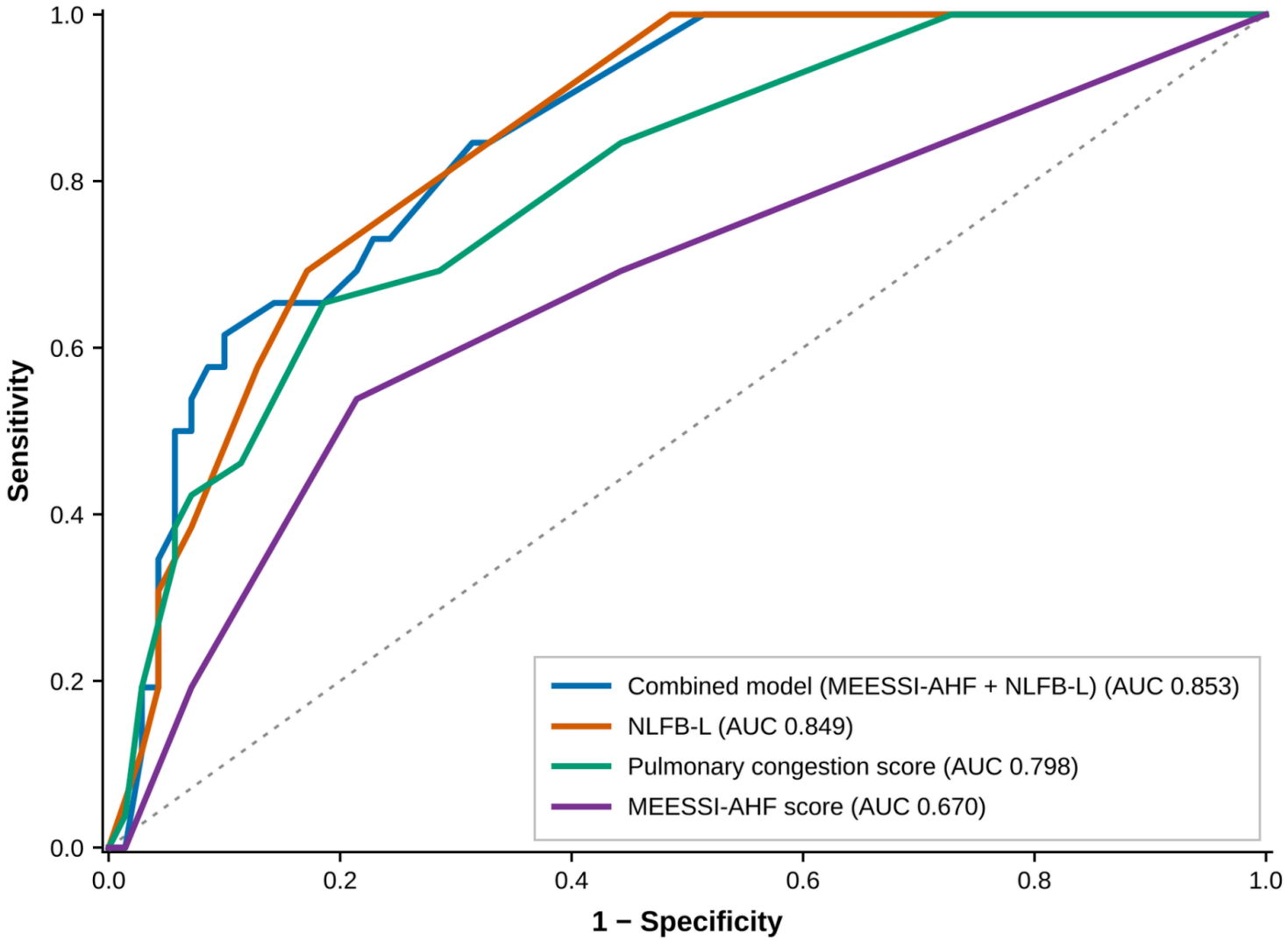


Figure 1

Receiver operating characteristic curves for the prediction of the 30-day primary outcome

Receiver operating characteristic (ROC) curves for the number of lung fields with B-lines (NLFB-L), the pulmonary congestion score and the MEESSI-AHF score individually, and for the combined model (NLFB-L plus MEESSI-AHF risk quintile), defined as the predicted probability from a logistic regression model including both variables. The area under the curve (AUC) of each model is shown in the key. The dotted diagonal line marks the line of no discrimination (AUC=0.50). The primary outcome was a 30-day heart failure-related emergency department revisit or hospital readmission. AUC, area under the ROC curve; MEESSI-AHF, Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure; NLFB-L, number of lung fields with B-lines.

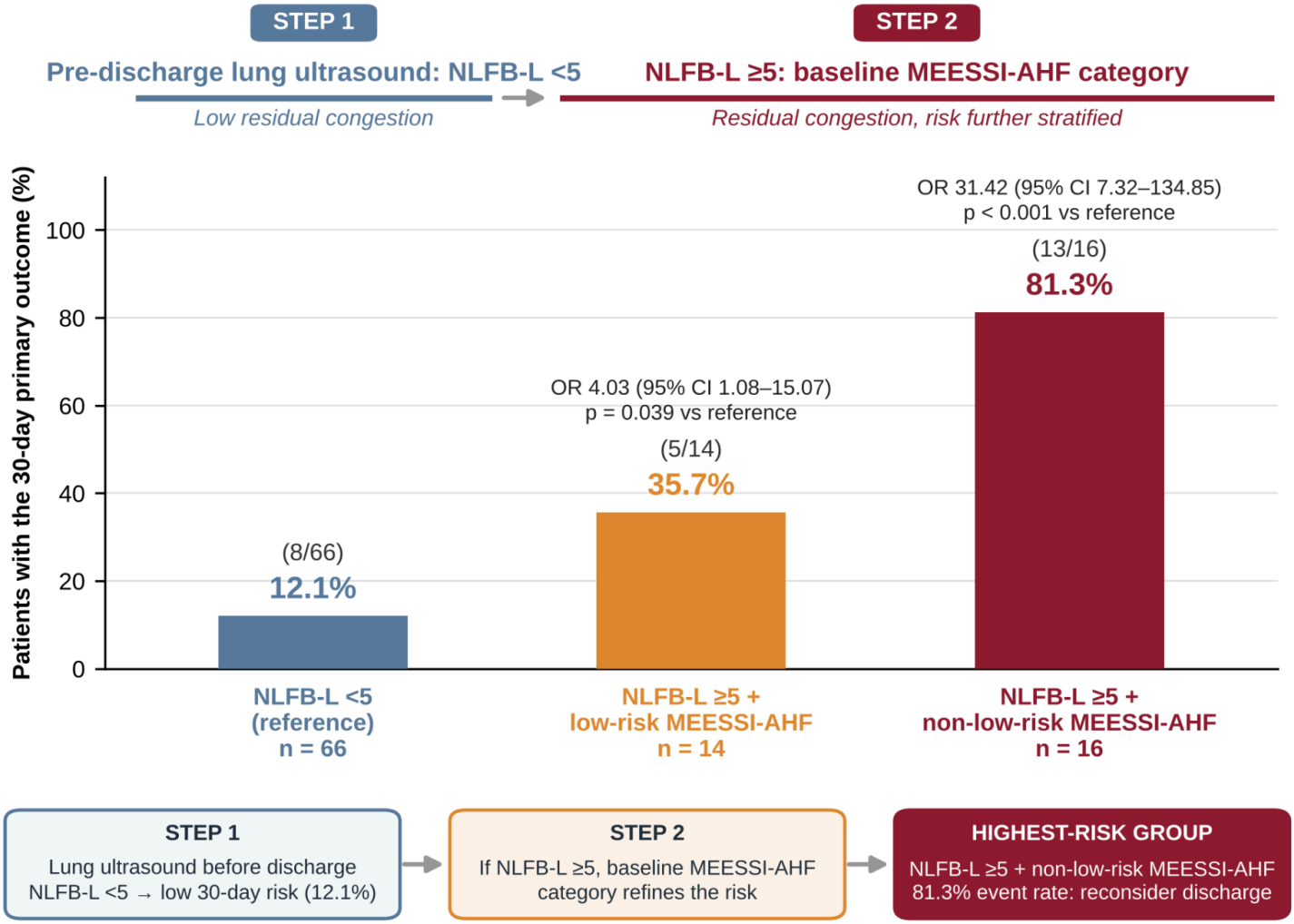


Figure 2

Observed 30-day event rates under the sequential echo-clinical strategy

Observed event rates under the proposed sequential echo-clinical strategy. Step 1: pre-discharge lung ultrasound assessed residual congestion; patients with NLFB-L <5 had a low event rate irrespective of baseline MEESSI-AHF risk and formed a single low-risk reference group. Step 2: among patients with NLFB-L ≥5, baseline MEESSI-AHF risk (low versus non-low) further stratified the primary outcome. The primary outcome was the composite of heart failure-related emergency department revisit or hospital readmission within 30 days. Bars show observed event rates with numerators and denominators; the odds ratios (OR) with 95% confidence intervals (CI) shown above the bars compare each NLFB-L ≥5 group with the reference group (NLFB-L <5) and were derived from the logistic regression model reported in Table 3. CI, confidence interval; MEESSI-AHF, Multiple Estimation of Risk based on the Emergency department Spanish Score in patients with Acute Heart Failure; NLFB-L, number of lung fields with B-lines; OR, odds ratio.

Supplementary Files

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

- Additionalfile1.docx
- Additionalfile2.docx
- Additionalfile3.xlsx