Clinical Policy: Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department With Acute Heart Failure Syndromes

This DRAFT is EMBARGOED – Not for Distribution

From the American College of Emergency Physicians Clinical Policies Subcommittee (Writing Committee) on Acute Heart Failure Syndromes:

Scott M. Silvers, MD (Subcommittee Chair)
Seth R. Gemme, MD
Sean Hickey, MD
Amal Mattu, MD
Jason S. Haukoos, MD, MSc (Methodologist)
Stephen J. Wolf, MD (Committee Co-Chair)

Members of the American College of Emergency Physicians Clinical Policies Committee (Oversight Committee):

Stephen J. Wolf, MD (Chair 2017-2021, Co-Chair 2021-2022)
Deborah B. Diercks, MD, MSc (Co-Chair 2021-2022)
Richard Byyny, MD, MSc (Methodologist)
Christopher R. Carpenter, MD, MSc
Seth R. Gemme, MD
Charles J. Gerardo, MD, MHS
Steven A. Godwin, MD
Sigrid A. Hahn, MD, MPH
Benjamin W. Hatten, MD, MPH
Jason S. Haukoos, MD, MSc (Methodologist)
Amy Kaji, MD, MPH, PhD (Methodologist)
Heemun Kwok, MD, MS (Methodologist)
Bruce M. Lo, MD, MBA, RDMS
Sharon E. Mace, MD
Devorah J. Nazarian, MD
Susan B. Promes, MD, MBA
Kaushal H. Shah, MD
Richard D. Shih, MD
Scott M. Silvers, MD
Andrea Slivinski, RN, DNP (ENA Representative 2021-2022)
Michael D. Smith, MD, MBA
Molly E. W. Thiessen, MD
Christian A. Tomaszewski, MD, MS, MBA
Jonathan H. Valente, MD
Melissa Villars, MD, MPH (EMRA Representative 2021-2022)
Stephen P. Wall, MD, MSc, MAEd (Methodologist)
Yanling Yu, PhD (Advocate for Patient Safety)
Stephen V. Cantrill, MD (Liaison with Quality and Patient Safety Committee)
John T. Finnell, MD (Board Liaison 2020-2022)
Travis Schulz, MLS, AHIP, Staff Liaison, Clinical Policies Committee, Clinical Policies Committee and Subcommittee on Acute Heart Failure Syndromes
Kaeli Vandertulip, MSLS, MBA, AHIP, Staff Liaison, Clinical Policies Committee

ABSTRACT

This clinical policy from the American College of Emergency Physicians is a revision of the 2007 *Clinical Policy: Critical Issues in the Evaluation and Management of Adult Patients Presenting to the Emergency Department with Acute Heart Failure Syndromes*. A writing subcommittee conducted a systematic review of the literature to derive evidence-based recommendations to answer the following clinical questions: 1) In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is the diagnostic accuracy of point-of-care lung ultrasound sufficient to direct clinical management? 2) In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is early administration of diuretics safe and effective? 3) In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is vasodilator therapy with high-dose nitroglycerin administration safe and effective? 4) In adult patients presenting to the emergency department with symptomatic acute heart failure syndrome, is there a defined group that may be discharged directly home for outpatient follow-up? Evidence was graded, and recommendations were made based on the strength of the available data.

INTRODUCTION

Heart failure continues to be a significant diagnosis that affects individuals in the United States at epidemic proportions, and the prevalence is growing. The prevalence of heart failure among adults in the United States has increased nearly 10% between 2012 (5.7 million Americans) and 2016 (6.2 million Americans). It is estimated that this prevalence will increase another 46% by 2030 to >8 million individuals.[1]

Acute heart failure Syndrome is a common condition encountered in the emergency department (ED), and it is associated with a 12% mortality rate during the in-hospital treatment period.[2] Although survival after the diagnosis of heart failure has improved slightly since 2012, 30-day, 1-year, and 5-year case fatality rates after hospitalization for heart failure remain high at 10%, 22%, and 42%, respectively.[1] The cost of this disease to the United States health system is high, and is expected to grow. In 2012, the total cost associated with heart failure was estimated to be $30 billion, and it is expected to increase to approximately $70 billion by 2030.[1]

The large heterogeneity of disease among acute heart failure patients has contributed to the reported definitions and terminology variability. As a result, it has been difficult to establish a consensus regarding the

EMBARGOED

actual definition, epidemiology, pathophysiology, and therapy for acute heart failure. The term "acute heart failure syndromes" emerged from the 2004 and 2005 meetings of an international workgroup convened primarily to establish uniform terminology and definitions in heart failure.[3,4] The workgroup defined acute heart failure syndromes as the "gradual or rapid deterioration in heart failure signs and symptoms resulting in a need for urgent therapy."[3] The consensus document further stated that these symptoms primarily manifest from increased pulmonary congestion that results from elevated left ventricular filling pressures (with or without low cardiac output) and may occur in patients with normal or reduced left ventricular ejection fraction.[3] Despite the need for standardization, these terms and definitions do not appear to have been more widely adopted in the literature since their initial publication.[3] Terms such as "acute decompensated heart failure" (ADHF) and "acute heart failure" (AHF) are still frequently used in the literature and for the purposes of this policy, are considered interchangeable. For consistency purposes, the subsequent discussion of individual studies in this policy will use the term acute heart failure syndrome (AHFS).

Appreciation of the heterogeneity in AHFS is important in the care of each individual patient. The ED plays a critical role in managing acute heart failure syndromes since approximately 80% of patients hospitalized for the condition are admitted through the ED. The comparison of studies to date has been made more challenging by the lack of consensus on what outcomes are most important (eg, cardiopulmonary indices, symptom relief, length of hospitalization, or morbidity and mortality).

This policy was intended to help improve the evaluation and management of heart failure patients presenting to an ED by answering 4 critical questions representing current interest or controversy.

METHODOLOGY

This ACEP clinical policy is based on a systematic review and critical descriptive analysis of the medical literature and is reported in accordance with PRISMA guidelines.

Search and Study Selection

This clinical policy is based on a systematic review with a critical analysis of the medical literature meeting the inclusion criteria. Searches of PubMed, SCOPUS, Embase, Web of Science, and the Cochrane Database of

EMBARGOED

Systematic Reviews were performed by a librarian. Search terms and strategies were peer-reviewed by a second librarian. All searches were limited to human studies published in English. Specific key words/phrases, years used in the searches, dates of searches, and study selection are identified under each critical question. In addition, relevant articles from the bibliographies of included studies and more recent articles identified by committee members and reviewers were included.

Two subcommittee members independently read the identified abstracts to assess them for possible inclusion. Of those identified for potential inclusion, each full-length text was reviewed for eligibility. Those identified as eligible were subsequently forwarded to the committee's methodology group (emergency physicians with specific research methodological expertise) for methodological grading using a Class of Evidence framework (Appendix A).

Assessment of Risk of Bias and Determination of Classes of Evidence

Each study identified as eligible by the subcommittee was independently graded by two methodologists. Grading was done with respect to the specific critical questions; thus, the Class of Evidence for any one study may vary according to the question for which it is being considered. For example, an article that is graded an "X" due to "inapplicability" for one critical question may be considered perfectly relevant for another question and graded I – III. As such, it was possible for a single article to receive a different Class of Evidence grade when addressing a different critical question.

Design 1 represents the strongest possible study design to answer the critical question, which relates to whether the focus was therapeutic, diagnostic, or prognostic, or a meta-analysis. Subsequent design types (ie, Design 2 and Design 3) represent respectively weaker study designs. Articles are then graded on dimensions related to the study's methodological features and execution, including but not limited to randomization processes, blinding, allocation concealment, methods of data collection, outcome measures and their assessment, selection and misclassification biases, sample size, generalizability, data management, analyses, congruence of results and conclusions, and potential for conflicts of interest.

Using a predetermined process that combines the study's design, methodological quality, and applicability to the critical question, two methodologists independently assigned a preliminary Class of Evidence grade for each

EMBARGOED

article. Articles with concordant grades from both methodologists received that grade as their final grade. Any discordance in the preliminary grades was adjudicated through discussion which involved at least one additional methodologist, resulting in a final Class of Evidence assignment (ie, Class I, Class II, Class III, or Class X) (Appendix B). Studies identified with significant methodologic limitations and/or ultimately determined to not be applicable to the critical question received a Class of Evidence grade "X" and were not used in formulating recommendations for this policy. However, content in these articles may have been used to formulate the background and to inform expert consensus in the absence of evidence. Question-specific Classes of Evidence grading may be found in the Evidentiary Table included at the end of this policy.

Translation of Classes of Evidence to Recommendation Levels

Based on the strength of evidence for each critical question, the subcommittee drafted the recommendations and supporting text synthesizing the evidence using the following guidelines:

Level A recommendations. Generally accepted principles for patient care that reflect a high degree of scientific certainty (eg, based on evidence from one or more Class of Evidence I, or multiple Class of Evidence II studies that demonstrate consistent effects or estimates).

Level B recommendations. Recommendations for patient care that may identify a particular strategy or range of strategies that reflect moderate scientific certainty (eg, based on evidence from one or more Class of Evidence II studies, or multiple Class of Evidence III studies that demonstrate consistent effects or estimates).

Level C recommendations. Recommendations for patient care that are based on evidence from Class of Evidence III studies or, in the absence of adequate published literature, based on expert consensus. In instances where consensus recommendations are made, "consensus" is placed in parentheses at the end of the recommendation.

There are certain circumstances in which the recommendations stemming from a body of evidence should not be rated as highly as the individual studies on which they are based. Factors such as consistency of results, uncertainty of effect magnitude, and publication bias, among others, might lead to a downgrading of recommendations. When possible, clinically-oriented statistics (eg, likelihood ratios [LRs], number needed to treat) are presented to help the reader better understand how the results may be applied to the individual patient. This can

EMBARGOED

assist the clinician in applying the recommendations to most patients but allow adjustment when applying to patients with extremes of risk (Appendix C).

Evaluation and Review of Recommendations

Once drafted, the policy was distributed for internal review (by members of the entire committee) followed by external expert review and an open comment period for all ACEP membership. Comments were received during a 60-day open comment period with notices of the comment period sent electronically to ACEP members, published in *EM Today*, posted on the ACEP Web site, and sent to other pertinent physician organizations. The responses were used to further refine and enhance this clinical policy, although responses do not imply endorsement. Clinical policies are scheduled for revision every 3 years; however, interim reviews are conducted when technology, methodology, or the practice environment changes significantly.

Application of the Policy

This policy is not intended to be a complete manual on the evaluation and management of adult patients with acute heart failure syndromes but rather a focused examination of critical questions that have particular relevance to the current practice of emergency medicine. Potential benefits and harms of implementing recommendations are briefly summarized within each critical question.

It is the goal of the Clinical Policies Committee to provide evidence-based recommendations when the scientific literature provides sufficient quality information to inform recommendations for a critical question. When the medical literature does not contain adequate empirical data to inform a critical question, the members of the Clinical Policies Committee believe that it is equally important to alert emergency physicians to this fact.

This clinical policy is not intended to represent a legal standard of care for emergency physicians. Recommendations offered in this policy are not intended to represent the only diagnostic or management options available to the emergency physician. ACEP recognizes the importance of the individual physician's judgment and patient preferences. This guideline provides clinical strategies for which medical literature exists to inform the critical questions addressed in this policy. ACEP funded this clinical policy.

EMBARGOED

Scope of Application. This guideline is intended for physicians working in EDs.

Inclusion Criteria. This guideline is intended for adult patients presenting to the ED with suspected acute heart failure syndrome.

Exclusion Criteria. This guideline is not intended for patients presenting with acute ST elevation myocardial infarction, high-output heart failure, cardiogenic shock, renal failure, valvular emergencies, pregnant patients, or pediatric patients.

CRITICAL QUESTIONS

1. In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is the diagnostic accuracy of point-of-care lung ultrasound sufficient to direct clinical management?

Patient Management Recommendations

Level A recommendations. None specified.

Level B recommendations. Use point-of-care lung ultrasound as an imaging modality in conjunction with history and physical examination to diagnose AHFS when diagnostic uncertainty exists.*

Level C recommendations. None specified.

* Use of lung ultrasound requires that the equipment is available, and the physician is proficient in its use.

Potential Benefit of Implementing the Recommendations:

- LUS provides greater diagnostic accuracy for AHFS than standard care.
- Improved time to diagnosis and treatment.

Potential Harm of Implementing the Recommendations:

- Lack of proficiency in lung ultrasound could lead to misdiagnosis.

Key words/phrases for literature searches: acute decompensated heart failure, acute diastolic heart failure, acute disease, acute heart failure, acute systolic heart failure, diagnostic imaging, echography, ED, emergencies, emergency, emergency care, emergency department, emergency health service, emergency medical services, emergency medicine, emergency room, emergency service, hospital emergency service, emergency services, emergency treatment, emergency ward, ER, heart failure, diastolic heart failure, systolic heart failure, hospital emergency service, lung, lung edema, lung POCUS, pulmonary ultrasonography, pulmonary US, ultrasonography, ultrasound, and variations and combinations of the key words/phrases. Searches included January 2007 to search dates of July 5, 2019, and June 17, 25, 26, and 29, 2020.

Study Selection: Five hundred three articles were identified in the searches. Sixteen articles were identified from the search results as candidates for further review. After grading for methodological rigor, zero Class I studies, 1 Class II study, and 8 Class III studies included for this critical question (Appendix D).

The use of lung ultrasound (LUS) to diagnose acute heart failure syndrome (AHFS) holds many benefits. LUS is a low-cost, rapid, non-ionizing imaging modality available at the bedside. LUS does not require an

EMBARGOED

inordinate amount of training or experience to become proficient and it has been demonstrated to be more accurate than a chest x-ray (CXR) in diagnosing pulmonary edema.[5,6] A diagnostic strategy incorporating bedside ultrasound has been shown to be superior in helping identify the correct diagnosis for undifferentiated dyspneic patients compared to a standard diagnostic strategy that did not incorporate ultrasound.[7] The use of bedside ultrasound specifically in the evaluation of AHFS is currently endorsed by the Society for Academic Emergency Medicine/Heart Failure Society of America's (SAEM/HFSA) Acute Heart Failure Working Group, and by the European Society of Cardiology's (ESC) consensus statement and heart failure guidelines (2015 and 2016).[4,8] This critical question evaluates the ability of LUS to accurately diagnose AHFS.

Diagnostic strategies that incorporate LUS have consistently been shown to be superior to evaluations without LUS in diagnosing AHFS.[9-17] A detailed review of the primary literature revealed a single Class II systematic review and metanalysis (SRMA), 3 Class III SRMAs, and 5 Class III studies that reported data pertinent to answering the critical question.

A 2018 Class II SRMA by McGivery et al[9] examined the accuracy of LUS in diagnosing AHFS among undifferentiated dyspneic ED patients. The systematic review included 7 studies and performed a metanalysis with a total sample of 1,861 patients. The pooled sensitivity and specificity for ED LUS for the diagnosis of AHFS was 82.5% and 83.6%, respectively, with a positive LR of 4.84 and a negative LR of 0.19. There was significant heterogeneity among the included studies. For this reason, a second metanalysis was performed, which included attending physicians only (excluded medical students and residents) and showed a sensitivity and specificity for ED LUS in the diagnosis of AHFS of 88.6% and 83.2%, respectively. Two studies included in this review found high inter-rater reliability when comparing novice sonographers to experts (k=82% and 92%, respectively). Further, one included study found that the LUS scans were completed in less than 1 minute while another found that it was completed in less than 5 minutes.

A Class III SRMA by Martindale et al[10] examined the diagnostic elements available to Emergency Physicians for the diagnosis of AHFS, including history and physical, electrocardiogram, CXR, natriuretic peptides, LUS, bedside echocardiogram, and bioimpedance. The diagnostic performance of LUS was shown to be superior to other diagnostic modalities. This SRMA included a total of 8 studies examining LUS for the diagnosis of AHFS in a total sample population of 1,918 patients. LUS was found to have a pooled sensitivity and

EMBARGOED

specificity of 85.3% and 92.7%, respectively, with a positive LR of 7.4 and a negative LR of 0.16. Comparatively, in this review, CXR was found to have a sensitivity of 56.9% and a positive LR of 4.8, and B-type natriuretic peptide (BNP) at a cutoff of <100 pg/ml was found to have a pooled sensitivity of 93.5%, a specificity of 52.9%, and a negative LR of 0.2. Bedside echocardiography identifying reduced ejection fraction was found to have a sensitivity and specificity of 80.6% for the diagnosis AHFS with a positive LR of 4.1 and a negative LR of 0.24.

A Class III SRMA by Staub et al[11] examined the accuracy of LUS in the diagnosis of AHFS, COPD/Asthma, and pneumonia. This SRMA included 14 studies with a total sample population of 2,778 patients where most patients were recruited from the ED. Overall the diagnostic accuracy of LUS for AFHS had an AUC of 0.91. This SRMA reported that the unpooled sensitivities and specificities for LUS ranged among studies from 75% to 90% for sensitivity and 80% to 90% for specificity. A second Class III SRMA by Lian et al[12] examined the accuracy of LUS for the diagnosis of AHFS in the ED. Fifteen studies were included with a total of 3,309 patients. The metanalysis found that the sensitivity and specificity were 85%, and 91%, respectively; the positive LR was 8.94, and the negative LR was 0.14. The AUC was 0.91. All 4 SRMAs' included the study by Pivetta et al[13] from 2015, which had a sample size of 1,005 patients. This study has also been reviewed separately as independent, primary literature. Table 1 summarizes the diagnostic performance of LUS for AHFS among the different meta-analyses.

Table 1. Summary of the Diagnostic Performance of Lung Ultrasound for Acute Heart Failure Syndrome as reported in 4 Meta-analyses

	SENS (95% CI)	SPEC (95% CI)	LR+ (95% CI)	LR- (95% CI)
McGivery et al[9] Class II N=1,861	82.5% (66.4% to 91.8%)	83.6% (72.4% to 90.8%)	4.84 (2.57 to 9.09)	0.19 (0.09 to 0.39)
Martindale et al[10] Class III N=1,918	85.3% (82.8% to 87.5%)	92.7% (90.9% to 94.3%)	7.4 (4.2 to 12.8)	0.16 (0.05 to 0.51)
Lian et al[12] Class III N=3,309	85% (84% to 87%)	91% (89% to 92%)	8.94 (5.64 to 14.18)	0.14 (0.08 to 0.26)
Staub et al[11] Class III N=2,778	(75% to 90%)	(80% to 90%)	NA	NA

Comparison of Sensitivity (SE), Specificity (SPE), Positive Predictive Values (PPV), Negative Predictive Values (NPV), Liklihood Ratios (LR) and Area under the Curve (AUC) of the included meta-analysis.

EMBARGOED

Three Class III studies directly compared the accuracy of LUS versus CXR with or without natriuretic peptides. In a 2019 diagnostic study of 518 patients by Pivetta et al[14] the authors compared the diagnostic accuracy of LUS versus CXR and natriuretic peptides in addition to clinical evaluation. This multi-centered, parallel randomized control trial included all adult patients presenting with a complaint of acute or acute on chronic dyspnea. After hospital discharge or death, 2 independent physicians (ie, an intensivist and emergency physician) reviewed the charts to adjudicate the etiology of the dyspnea. The accuracy of the diagnosis of AHFS by clinical examination alone did not differ between the 2 arms. However, LUS was found to be more accurate than clinical evaluation alone and more accurate than the combination of clinical exam, CXR, and natriuretic peptides. Also notable in this study is the fact that the strategy of CXR combined with natriuretic peptides did not significantly increase the diagnostic accuracy compared to clinical evaluation alone. The authors concluded that the approach utilizing LUS reduced diagnostic errors in 8% of patients and it also reduced the median time to diagnosis from 104.5 minutes to 5 minutes.

In an earlier, prospective, multicenter, Class III study of 1,005 patients, Pivetta et al[13] (2015) reported improved diagnosis of AHFS using LUS. In this study of adult patients with acute or acute on chronic dyspnea, patients had a standard work up followed by questioning of the examining physician as to whether they believed that the cause of the dyspnea was due to AHFS. LUS was then performed, and the presumptive diagnosis was reassessed. Following discharge, the patients' final diagnosis was adjudicated by an emergency physician and a cardiologist, both of whom were blinded to the LUS results. Standard clinical work up was shown to be inferior compared to a diagnostic strategy that incorporated LUS for the diagnosis of AHFS (Table 2). LUS alone was also shown to be superior to both CXR as well as natriuretic peptides (ie, BNP/NT-prop-BNP) in the diagnosis of AHFS.

This study was followed by another Class III study by Sartini et al.[15] In this prospective single-center observational cohort study of 236 adult patients with acute or acute on chronic dyspnea, investigators examined the diagnostic accuracy of LUS, CXR, and NT-pro-BNP in the diagnosis of AHFS. Emergency physicians skilled in LUS performed the examinations and were blinded to all other aspects of patient care. The sensitivity of LUS reported in this study of 57.73% was lower than that reported in other studies. However, a subgroup analysis of LUS performance among patients who did not receive pre-hospital diuretics found that the sensitivity of LUS was

EMBARGOED

83%, which is consistent with the other studies. The transport times were not disclosed in the study, so it is difficult to assess how likely it was for the administration of a diuretic to affect the findings on LUS.

EMBARGOED

313 **Table 2.** Standard clinical workup versus diagnostic strategy with LUS to diagnose AHFS.

	Design		SE (95% CI)	SPE (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR- (95% CI)	AUC (95% CI)
Pavetta et al 2019[14] N=518 Class III	Multi-centered Randomized Controlled Trial Study	LUS	93.5 (87.7 to 97.2)	95.5 (90.5 to 98.3)	95.1 (89.6 to 98.2)	94.1 (88.7 to 97.4)	20.9 (9.54 to 45.7)	0.07 (0.03 to 0.13)	0.95
		CXR/NT-pro-BNP	85 (76.5 to 91.4)	89.4 (83.5 to 93.7)	83.3 (74.7 to 90.0)	90.5 (84.8 to 94.6)	8.0 (5.1 to 12.6)	0.17 (0.11 to 0.27)	0.87
Pavetta et al 2015[13] N=1,005 Class III	Prospective Multi-Centered Observational Cohort study	Clinical Work-up	85.3% (81.8 to 88.4)	90% (87.2 to 92.4)	88% (84.6 to 90.8)	87.8% (84.8 to 90.4)	8.6	0.2	87.6
		LUS-Implemented	97% (95 to 98.3)	97.4 % (95.7 to 98.6)	97% (95 to 98.3)	97.4% (95.7 to 98.6)	37.5	0.03	0.97
		LUS-Alone	90.5% (87.4 to 93)	93.5 (91.1 to 95.5)	92.3% (89.4 to 94.6)	92% (89.4 to 94.1)	14	0.1	0.92
		CXR	69.5% (65.1 to 73.7)	82.1 (78.6 to 85.2)	76.8% (72.5 to 80.8)	75.9% (72.5 to 79.3)	3.9	0.4	0.76
Pavetta et al 2015[13] Subgroup analysis N=486		LUS-Implemented	97.5% (94.9 to 99)	95.6% (91.9 to 98)	96.8% (94 to 98.5)	96.6% (93.1 to 98.6)	22.3	0.02	0.97
		BNP/NT-pro-BNP	85% (80.3 to 89)	61.7% (54.6 to 68.3)	75.1% (69.9 to 79.7)	75.1% (67.9 to 81.6)	2.2	0.20	0.73
		LUS-alone	89.3% (85.1 to 92.7)	89.8% (84.8 to 93.6)	92.3% (88.4 to 95.1)	86% (80.7 to 90.4)	8.8	0.11	0.90
Sartini et al 2017[15] N=236 Class III	Prospective Single Centered Observational Cohort Study	LUS	57.73% (47.28 to 67.7)	87.97% (81.2 to 92.96)	77.78% (66.4 to 86.73)	74.05% (66.49 to 80.69)	4.8 (2.94 to 7.83)	0.48 (0.38 to 0.61)	0.84
		CXR	74.49% (64.69 to 82.76)	86.26% (79.16 to 91.65)	80.22% (70.55 to 87.84)	81.88% (74.43 to 87.92)	5.42 (3.48 to 8.45)	0.30 (0.21 to 0.42)	
		NT-pro-BNP >300 pg/ml	97.59% (91.57 to 99.71)	27.56% (20.01 to 36.19)	46.82% (39.21 to 54.54)	94.59% (81.81 to 99.34)	1.35 (0.20 to 1.51)	0.09 (0.02 to 0.35)	0.76
		LUS subgroup without pre-hospital diuretics N=181	83%	86.39%	N/A	N/A	N/A	N/A	N/A

EMBARGOED

314 Comparison of Sensitivity (SE), Specificity (SPE), Positive Predictive Values (PPV), Negative Predictive Values (NPV), Likelihood Ratios (LR) and Area under
315 the Curve (AUC) of the included studies.

EMBARGOED

A Class III single-center observational cohort in Thailand by Nakornchai et al[16] assessed whether multi-organ point-of-care ultrasound (POCUS) performed by emergency medicine residents could be used to improve the diagnostic accuracy of AHFS. This study has major limitations and excludes patients with myocardial infarction, shock, or those receiving positive pressure ventilation. Furthermore, its main outcome was the diagnostic accuracy of multi-organ POCUS instead of solely examining the diagnostic accuracy of LUS for AHFS. These limitations could explain the uniquely poor sensitivity (35%) reported for LUS in diagnosing AHFS in this study compared to the other examined literature, though the specificity (72.7%) reported is consistent with that of other examined literature.

A Class III, multi-center, prospective, observational cohort study by Buessler et al[17] examined the use of LUS in addition to the BREST score (clinical decision tool for diagnosing heart failure) in diagnosing AHFS. This study found that LUS increased diagnostic accuracy in addition to the BREST score both in the whole patient population and in patients with intermediate BREST scores.

Summary

In patients presenting with acute dyspnea and the possible diagnosis of AHFS, evidence supports the use of POCUS to improve diagnostic accuracy and help direct management. The presence of B-lines on bedside ultrasound is an independent predictor of AHFS. When combined with historical information and physical examination findings, bedside ultrasound outperforms chest radiography and laboratory testing including natriuretic peptides.

Future Research

To date, no studies have evaluated if the more rapid diagnosis of acute decompensated heart failure using LUS significantly alters important clinical patient-centered outcomes. A randomized control trial that compares the use of LUS to identify B-lines versus usual care on outcomes such as need for intubation, intensive care unit (ICU) admissions, and mortality would be the next logical step. Additionally, randomized control trials are also needed to examine whether the use of a multi-modal POCUS strategy significantly improves the standard diagnostic work-up for patients being considered for the diagnosis of AHFS in the ED.[18,19]

EMBARGOED

2. In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is early administration of diuretics safe and effective?

Patient Management Recommendations

Level A recommendations. None specified.

Level B recommendations. None specified.

Level C recommendations. Although no specific timing of diuretic therapy can be recommended, physicians may consider earlier administration of diuretics, when indicated for ED patients with AHFS, as it may be associated with reduced length of stay and in-hospital mortality (Consensus recommendation).

Physicians should be confident in the diagnosis of acute heart failure syndrome with volume overload in a patient prior to the administration of diuretics as treatment with diuretics may cause harm to those with an alternative diagnosis (Consensus recommendation).

Potential Benefit of Implementing the Recommendations:

- Decrease delays in treatment of concomitant conditions.
- Decrease length of stay and inpatient mortality.

Potential Harm of Implementing the Recommendations:

- Giving diuretics too early to a patient who is ultimately proven not to have the diagnosis of AHFS or when the patient is not experiencing volume overload as a cause of their AHFS could be harmful.

Key words/phrases for literature searches: acute decompensated heart failure, acute diastolic heart failure, acute heart failure, acute systolic heart failure, diuretic, diuretic agent, diuretics, ED, emergencies, emergency care, emergency department, emergency health service, emergency medical services, emergency medicine, emergency room, emergency service, emergency services, emergency treatment, emergency ward, ER, heart failure, hospital emergency service. Searches included January 2007 to search dates of July 8, 2019, and June 22, 25, and 29, 2020.

Study Selection: Five hundred eighty-three articles were identified in the searches. Eleven articles were identified from the search results as candidates for further review. After grading for methodological rigor, zero Class I studies, zero Class II studies, and 1 Class III study was included for this critical question (Appendix D).

The use of loop diuretics in the management of acute heart failure induces an increase in sodium and water excretion by the kidney, thus reducing preload on the heart. It has been an integral component of the multimodal management of acute heart failure patients with volume overload in the ED for the last 40 years. The

EMBARGOED

management of patients with euvolemic or hypovolemic heart failure is more complex and typically requires alternative therapeutic strategies; however, this is beyond the scope of the critical question reviewed. When a patient with heart failure is found to be volume overloaded, the loop diuretics that are frequently administered are furosemide, torsemide, and bumetanide. The pharmacodynamics differ between these medications when given intravenously, with furosemide having the least potency followed by torsemide and bumetanide. Torsemide and bumetanide have similar bioavailability, which is higher than that of furosemide. The time to peak effect for these medications ranges from 15 minutes for bumetanide to 30 minutes for furosemide, and 60 minutes for torsemide. Despite the known pharmacodynamics of these medications, their optimal timing of administration in the ED and the subsequent effect on clinical outcomes are unclear.

Defining "early" treatment in the ED is difficult due to various definitions in the literature. There has not been a widely accepted timing administration goal with regards to diuretics in AHFS as there has been with therapies for other disease processes. Further confounding the "early administration" of diuretics is the fact that heart failure is not easy to quickly identify among undifferentiated patients in the ED. Some patients ultimately require admission to the hospital and further studies such as echocardiography before a more definitive diagnosis can be made. The majority of patients who receive "early" identification and treatment tend to be those who have had previous episodes of established AFHS and develop similar symptoms or those with more severe and classic symptoms. Therefore, with regards to our search for this question, we purposely did not define early by a specific time cutoff, but rather left it undefined to help ensure that we captured all applicable literature. Unfortunately, many of the studies that have addressed the question of the timing of administration have been of limited quality and/or not applicable to the US ED population.

In a Class III, observational trial by Wong et al,[20] authors did not find an association between treatment delays and 30-day all-cause mortality or readmission. Although, they did find an association between treatment delays and other outcomes. This study was a retrospective secondary analysis of 6,971 patients from the Acute Decompensated Heart Failure Registry Emergency Module (ADHERE-EM). Patients ≥65 years old who were hospitalized for AHFS and received intravenous heart failure therapy at the initial visit were studied. The median time to IV heart failure therapy was 2.3 hours (1.1 to 4.4 hours), with an incidence of 30-day all-cause mortality or readmission of 27.4%. Time to treatment had no clinically significant association with 30-day all-cause

EMBARGOED

mortality (hazard ratio (HR) 1.00; 95% CI 1.00 to 1.01) or to 30-day all-cause readmission (HR 1.00; 95% HR 0.99 to 1.00). Increasing time to treatment was associated with a very small increased risk of in-hospital mortality (HR 1.01; 95% CI 1.00 to 1.02), as well as an approximate 1.4 hour increase in index admission length of stay. This included treatments with a diuretic alone and combinations of a diuretic with an inotrope, or vasoactive medication. However, it is important to note that these statistically significant results are unlikely clinically significant. They did perform a subset cohort analysis, including those patients who did receive a diuretic or a diuretic and another agent. In those patients receiving diuretics alone, there was also no difference in all-cause mortality at 30 days.

Summary

Only one weaker, class III study was identified that met criteria and helped to answer this critical question. Therefore, it is difficult to make confident recommendations related to the timing of diuretic therapy in patients with AHFS. The decision to treat early is complicated by the fact that rapidly and accurately identifying AHFS is often difficult, and administration of diuretics to patients without volume overload and the diagnosis of acute heart failure may cause harm.

Future Research

Future research should involve randomization of patients presenting to the ED with suspected acute heart failure to treatment with intravenous diuretics at clearly defined time intervals with clinically significant outcome measures such as hypotension, kidney injury, need for escalation of therapy or level of monitoring, length of stay, and mortality. Research should also focus on factors that help to accurately and rapidly identify the diagnosis of AHFS.

3. In adult patients presenting to the emergency department with suspected acute heart failure syndrome, is vasodilator therapy with high-dose nitroglycerin administration safe and effective?

Patient Management Recommendations

Level A recommendations. None specified.

Level B recommendations. None specified.

EMBARGOED

Level C recommendations. Consider using high-dose nitroglycerin as a safe and effective treatment option when administered to patients with AHFS and elevated blood pressure (Consensus Recommendation).*

* While nitroglycerin infusions up to 400 mcg/min have been described as "standard dosing"[21] some may consider dosing of 200 mcg/min or higher also as "high dose." "High dose" nitroglycerin has also been described as bolus intravenous dosing of 2000 mcg every 3 to 5 minutes.[21]

Potential Benefit of Implementing the Recommendations:

- This therapy has the potential benefit of reducing respiratory distress and decreasing the need for endotracheal intubation and ICU admission.

Potential Harm of Implementing the Recommendations:

- High-dose nitrates could potentially produce hypotension.

Key words/phrases for literature searches: acute decompensated heart failure, acute diastolic heart failure, acute disease, acute heart failure, acute systolic heart failure, ED, emergencies, emergency, emergency care, emergency department, emergency health service, emergency health services, emergency medicine, emergency room, emergency service, emergency services, emergency treatment, emergency ward, ER, glyceryl triturate, heart failure, hospital emergency service, nitroglycerin. Searches included January 2007 to search dates of July 12 and 15, 2019, and June 23, 25, and 29, 2020.

Study Selection: One hundred seventy-seven articles were identified in the searches. Fifteen articles were identified from the search results as candidates for further review. After grading for methodological rigor, zero Class I studies, zero Class II studies, and 1 Class III study was included for this critical question (Appendix D).

Our knowledge of the pathophysiology of AHFS has evolved over the years. Whereas it was once thought that AHFS was simply a problem of volume overload, we now know that many cases are the result of cardiac dysfunction and excess vascular tone.[22] The majority of patients presenting with AHFS suffer from excess preload and an increased systemic vascular resistance (afterload), which is associated with a reduced cardiac output via systolic or diastolic dysfunction.[21] Vasodilators can reduce preload and afterload, resulting in improved cardiac output and reduced pulmonary congestion. The use of vasodilators has therefore been recommended in consensus statements for the emergent management of AHFS.[22,23]

Nitrates have been the preferred vasodilator in the management of AHFS for decades and are part of the standard recommended medication regimen for patients who are normotensive or hypertensive.[22,23] Nitrates decrease preload and, at higher doses, decrease afterload as well. The onset of vasodilatory effects is within 1 to 3 minutes, depending on the route of delivery, and the half-life is 2 to7 minutes. Unfortunately, consensus statements and guidelines provide little direction as to the optimal dosing regimen of nitrates. The American

EMBARGOED

College of Emergency Physicians' prior clinical policy pertaining to the evaluation and management patients presenting to the ED with AHFS[3] addressed whether vasodilator therapy should be prescribed in the ED for the patient with AHFS. The 2007 Clinical Policy concluded with a Level B recommendation that intravenous nitrate therapy should be used; however, no specific dosing was recommended.

In the only Class III trial of the above studies, Levy et al[21] performed a nonrandomized open-label trial evaluating the use of 2000 mcg intravenous bolus doses of nitroglycerin every 3 to 5 minutes in patients with dyspnea and AHFS whose systolic blood pressure was $\geq$160 mm Hg or whose mean arterial pressure was $\geq$120 mm Hg and who were refractory to initial therapy. Initial therapy consisted of Class I recommendations from the American Heart Association's 2000 "Guidelines for the Evaluation and Management of Heart Failure,"[24] including high-flow oxygen, with 100% nonrebreather mask, sublingual nitroglycerin 400 mcg every 5 minutes up to a maximum of 4 treatments, and furosemide 60 to 80 mg IV. Morphine sulfate 3 to 5 mg intravenous push was considered optional. If patients failed to improve to this initial therapy and the treating physician believed that the patient required intravenous nitroglycerin, consenting patients were included in the trial.

Trial patients received an intravenous infusion of nitroglycerin started at a dose of 0.3 to 0.5 mcg/kg/min. The nitroglycerin infusion was increased at the discretion of the treating physician in increments of 20 mcg/min every 1 to 3 minutes to a maximum of 400 mcg/min. If the systolic blood pressure fell below 90 mm Hg, further increased dosing of the nitroglycerin was discontinued. If blood pressure did not improve, the nitroglycerin was stopped. Concurrent with the initiation of the nitroglycerin infusion and titration, all patients received an initial 2000 mcg intravenous bolus of nitroglycerin, and repeat dosing of the 2000 mcg intravenous boluses was allowed every 3 to 5 minutes at the discretion of the treating physician for a period of up to 30 minutes (maximum potential dose of 20 mg). The mean total dose of bolus-dose nitroglycerin was 6500 mcg ($\pm$3400 mcg).

The patients receiving high-dose nitroglycerin showed a trend toward improvement in the primary efficacy endpoint of reduced endotracheal intubations within 6 hours [13.8% (95% CI 4.8% to 29.5%) versus 26.7% (95% CI 15.5% to 40.8%)]; a trend toward improvement in the secondary endpoints of reduced need for bilevel positive airway pressure (BiPAP) [6.9% (95% CI 1.5% to 20.3%) versus 20.0% (95% CI 10.4% to 33.3%)] and an improvement in the secondary endpoint of reduced intensive care unit admissions (37.9% (95% CI 22.1% to 56.1%) versus 80.0% (95% CI 66.7% to 89.6%)]. A single episode of symptomatic hypotension

EMBARGOED

occurred in the high-dose nitroglycerin group after a single bolus of 2 mg but resolved after a 500 ml intravenous bolus of fluid. No patient in either group developed immediate electrocardiogram (ECG) changes of ischemia. Neither the high-dose nitroglycerin group nor the nonintervention group demonstrated any adverse neurologic events or in-hospital deaths.

Summary

Acute heart failure syndrome is a common ED presentation, often associated with ICU admission and endotracheal intubation. Guidelines and consensus statements now recognize the importance of cardiac and vascular dysfunction in the pathophysiology of AHFS, but specific recommendations regarding vasodilator therapy (medication, dosing regimen, etc.) are lacking. The single Class III study noted above demonstrated the safety of high-dose nitroglycerin therapy in patients with AHFS while suggesting possible improvements in clinical outcomes, including reduced need for endotracheal intubation, BiPAP, and ICU admission without significant adverse effects.

Future Research

Given the frequency of AHFS presentations to the ED and its associated morbidity and mortality, it is critical that research continue to inform physicians about the optimal care of these patients. The current prevailing theory regarding the pathophysiology underlying many of these presentations focuses on excess preload as well as vascular tone dysfunction manifesting as excess afterload. Studies evaluating the use of high-dose nitrates to treat this vascular tone dysfunction have been favorable in their trend towards a larger magnitude of benefit; however, current studies are limited by their small numbers and their retrospective, nonrandomized, open-label designs. Larger studies utilizing a prospective randomized, blinded protocol would be invaluable in clarifying whether high-dose nitrates can, in fact, produce rapid clinical benefits in selected patients with AHFS.

4. In adult patients presenting to the emergency department with symptomatic acute heart failure syndrome, is there a defined group that may be safely discharged home for outpatient follow-up?

Patient Management Recommendations

EMBARGOED

Level A recommendations. None specified.

Level B recommendations. Do not rely on current AHFS risk stratification tools alone to determine which patients may be discharged directly home from the ED.

Consider using the Ottawa Heart Failure Risk Scale (OHFRS) to help determine which higher-risk patients for the adverse outcome should not be discharged home.

Level C recommendations. Consider using the Emergency Heart Failure Mortality Risk Grade (EHMRG7) or the STRATIFY Decision Tool to help determine which higher-risk patients for the adverse outcome should not be discharged home.

Use shared decision-making strategies when determining the appropriate disposition of AHFS patients (Consensus recommendation).

Potential Benefit of Implementing the Recommendations:

- ED physicians may reduce the likelihood that a discharged patient experiences an adverse outcome during short-term follow-up.

Potential Harm of Implementing the Recommendations:

- ED physicians may increase the number of AHFS admissions to the hospital, which would potentially add to hospital overcrowding and negatively impact reported AHFS readmission metrics.
- More patients could experience complications associated with hospital admission.

Key words/phrases for literature searches: acute decompensated heart failure, acute diastolic heart failure, acute disease, acute heart failure, acute systolic heart failure, discharge, discharged, ED, emergencies, emergency, emergency care, emergency department, emergency health service, emergency medical services, emergency medicine, emergency room, emergency service, emergency services, emergency treatment, emergency ward, ER, heart failure, hospital discharge, hospital emergency service, patient discharge. Searches included January 2007 to search dates of July 15, 2019, and June 25 and 29, 2020.

Study Selection: Nine hundred thirty-eight articles were identified in the searches. Fourteen articles were identified from the search results as candidates for further review. After grading for methodological rigor, zero Class I studies, 1 Class II studies, and 3 Class III studies were included for this critical question (Appendix D).

Hospital admissions account for an estimated one-third of healthcare spending in the United States.[25] Approximately one-half of all hospital admissions in the United States originate from EDs.[25] Of the ~$39.2 billion dollars spent on heart failure care in the United States each year, hospital admissions account for a total cost of over $11 billion, and they represent the single largest proportion of the expenditure.[26,27] Patients with heart failure

EMBARGOED

578 often have other co-morbid conditions (eg, hypertension, ischemic heart disease, diabetes, chronic kidney disease
579 and atrial fibrillation), which may also contribute to admission decisions and prolong hospital length of stay.[27,28]
580 Heart failure is a relatively grave diagnosis as it is associated with high 30-day, 1 year, and after hospitalization of
581 10%, 22%, and 42%, respectively.[29] Although more than 80% of all ED visits in the United States result in the
582 patient being discharged home, 82% of patients presenting with primary AHFS are admitted to the hospital for
583 further care with a median inpatient length of stay of 3.4 days.[27,30] This differs from other countries such as
584 Canada where the admission rate of acute heart failure patients from the ED is estimated to be only 40 to 60%.[31]
585 Patients admitted for heart failure have readmission rates as high as 30% to 60% within 3 to 6 months after
586 discharge.[28]

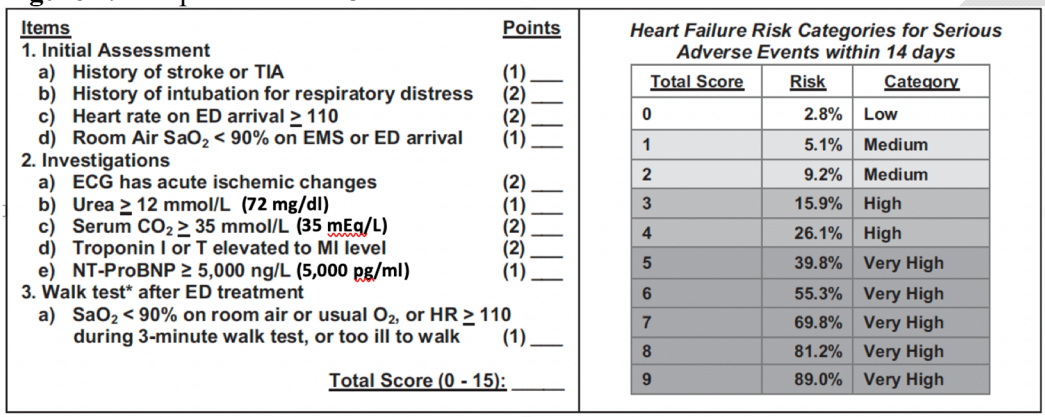
587 Although over 80% of all patients presenting to an ED in the United States with primary AHFS are
588 admitted to the hospital, approximately one-half present with "low risk" features and are believed to be possibly
589 unnecessary admissions.[28,30] The historical reluctance of emergency medicine physicians to discharge a greater
590 percentage of acute heart failure patients home could be the result of several factors including the significant
591 mortality associated with the disease, the relatively high associated complication rate including readmissions, and
592 the absence of any known discrete risk factors or decision rules that could help reliably establish which patients
593 are safe to be discharged directly home from the ED. Establishing a low-risk group of AHFS patients who were
594 safe for direct discharge from the ED could reduce healthcare costs, reduce the risk of nosocomial infections and
595 other untoward events associated with hospital stays, improve the availability of hospital beds for sicker patients,
596 and improve patient satisfaction.

597 Therefore, a comprehensive review of the medical literature was performed to learn if any data could be
598 used to reliably define which patients, presenting to an ED with the diagnosis of symptomatic AHFS, could be
599 safely discharged directly home. The literature review revealed 56 publications that were deemed potentially
600 applicable to the critical question. After further analysis, 42 of these articles were assessed as not directly
601 addressing the critical question. Fourteen studies were identified as pertinent, reviewed by the methodologists,
602 and received grading. Of these 14 studies, 10 were considered fatally flawed, and 4 studies (1 class II, and 3 class
603 III) were considered contributory and received a grade.

EMBARGOED

In the only class II study referred to as "RAD-2", Stiell et al prospectively studied consecutive adult patients who presented to the ED with serious adverse events (SAEs) resulting from acute heart failure, to assess the accuracy, acceptability, and potential impact of their previously derived Ottawa Heart Failure Risk Scale (OHFRS) score (from the RAD-1 Study, see Figure 1) on a new population of patients.[31,32]

Figure 1. Comparison of the 3 Risk Stratification Tools.[31]



Items	Points
1. Initial Assessment	
a) History of stroke or TIA	(1) ___
b) History of intubation for respiratory distress	(2) ___
c) Heart rate on ED arrival ≥ 110	(2) ___
d) Room Air SaO_2 < 90% on EMS or ED arrival	(1) ___
2. Investigations	
a) ECG has acute ischemic changes	(2) ___
b) Urea ≥ 12 mmol/L (72 mg/dl)	(1) ___
c) Serum CO_2 ≥ 35 mmol/L (35 mEq/L)	(2) ___
d) Troponin I or T elevated to MI level	(2) ___
e) NT-ProBNP ≥ 5,000 ng/L (5,000 pg/ml)	(1) ___
3. Walk test* after ED treatment	
a) SaO_2 < 90% on room air or usual O_2, or HR ≥ 110 during 3-minute walk test, or too ill to walk	(1) ___
Total Score (0 - 15):	____

Heart Failure Risk Categories for Serious Adverse Events within 14 days

Total Score	Risk	Category
0	2.8%	Low
1	5.1%	Medium
2	9.2%	Medium
3	15.9%	High
4	26.1%	High
5	39.8%	Very High
6	55.3%	Very High
7	69.8%	Very High
8	81.2%	Very High
9	89.0%	Very High

(PENDING COPYRIGHT PERMISSION)

The OHFRS score was developed to estimate the probability of SAEs in the subsequent 14 days after evaluation and treatment in the ED. In this study, the use of the OHFRS score at different cutoff points was unable to identify a group of patients who were reliably safe for discharge directly home. However, the score did perform better than standard physician decision-making in predicting which patients should not be discharged home due to their higher risk of serious adverse outcomes.

In this study, 4,999 patients presenting to the ED with shortness of breath were screened for eligibility. Of these, 3,130 were deemed ineligible and excluded from the study. Patients were excluded from the study if they were believed to be unstable or subjectively "too ill" to be considered for discharge after 2 to 12 hours of ED management. Of the 1,869 eligible patients, 769 were missed, primarily for presenting outside of the study hours, and 1,100 were enrolled in the study. Included patients had an OHFRS score calculated 2 to 12 hours after ED treatment. After calculating the OHFRS score, staff were asked which risk category the patient was in (ie, low, medium, high, very high) for a serious adverse event, and how comfortable they would be to use the scale to make a disposition decision (5-point scale from very comfortable to very uncomfortable).

EMBARGOED

The primary outcome was any serious adverse event defined as: death from any cause within 30 days of the ED visit, any of the following within 14 days of the visit [admission to a monitored unit, endotracheal intubation or need for noninvasive ventilation (unless used at home), myocardial infarction, recipient of a major procedure (eg, percutaneous coronary intervention, cardiac surgery, or hemodialysis)], or returning to the ED for any medical problem related to the initial presentation within 14 days resulting in admission to the hospital. Investigators assessing for SAEs were blinded to the OHFRS scores. Of the 1,100 enrolled patients, 684 received NT-proBNP testing (a component of the score) at their index evaluation.

Using their normal standard of care, researchers admitted 57.2% of patients to the hospital, and discharged 42.8% from the ED at the index visit. The overall SAE rate was 15.5% with 19.4% occurring among admitted patients and 10.2% among those discharged home. Of the 41 recorded deaths, 16 (39%) occurred among those patients who were discharged home. Using an OHFRS score cutoff of >1 among patients both without the NT-proBNP value as well as those with the value would have increased the sensitivity for the outcome of an SAE from approximately 70% using clinical judgment alone to 91.8% and 95.8%, respectively (still missing as many as 8% and as few as 4% of SAEs). However, this increased sensitivity would have also led to a 20% to 26% respective increase in the admission rate of the two groups of patients. Overall, the researchers reported that 11.9% of the time they felt "uncomfortable" or "very uncomfortable" in using the OHFRS to make disposition decisions for their patients.

Although this study may be the highest quality of the studies to date looking prospectively at predictors of outcomes among ED patients with AHFS, it still has several limitations including intention bias as patients who were admitted to the hospital may be less likely to experience an SAE due to the closer monitoring, the fact that not all patients had NT-pro-BNP measured, not all patients received assessment while ambulating, enrolled subjects were, for the most part, a convenience sample as some patients were not included due to the researchers being "too busy," and the fact that the study used the same academic EDs that derived the original OHFRS score. Although never mentioned, it is assumed that when NT-proBNP was not measured, a score of "0" was used for this variable. Additionally, patients who were believed to be subjectively "too ill to be ready for discharge after 2 to 12 hours of ED management" were excluded, so we don't fully understand how the OHFRS would have performed if applied to all presenting AHFS patients. Also of note, adding a point on the scale when patients were

EMBARGOED

too sick or never asked to do the walk test, as the investigators did, could falsely risk stratify the patient to a group of less risk than had the variable actually been assessed. Finally, since the study discharged patients home based on the standard of physician gestalt, there is still no reliable data as to how the OHFRS score would perform for discharge decision-making if it were the sole determinant for disposition.

Given the current data including both the seriousness of the adverse outcomes assessed and the limitations of the study, the OHFRS score appears neither sensitive nor specific enough to be relied upon as the sole criteria by which ED physicians could make disposition decision-making for their AHFS patients. The sensitivity of the tool when all OHFRS variables are recorded does, however, hold promise for future study. It would be interesting to learn how the scale could impact care among AHFS patients cared for by healthcare professionals in United States EDs where admission rates are significantly higher. Still, knowledge of the different OHFRS variables and the different variables' performance in screening for the likelihood of an SAE could further aid ED physicians in the disposition decision-making of their AHFS patients.

Although Stiell's 2017 report of the OHFRS score, "RAD-2,"[31] is a better validated, prospectively studied report of the use of the OHFRS score, his original derivation class III study of the score in 2013, "RAD-1,"[32] is notable for several reasons. First, the rate of admission for the AHFS patients from the ED in RAD-1 was much lower 38.1% in the original study (versus 57.2% in RAD-2), and the rates of SAEs were also slightly lower for both admitted patients 16% (versus 19.4% in RAD-2) and discharged patients 9.0% (versus 10.2% in RAD-2). Despite these facts, the diagnostic performance for the score remained relatively consistent between RAD-1 and RAD-2 datasets for both sensitivity and specificity at the different cutoffs of >1 and >2 points. This confirmation of the association between the score and the outcome of SAEs is reassuring that future applications of the score, at least in that region of hospitals, will likely yield similar results.

Of the 2 remaining class III studies, only Collins et al[33] in 2015 reported meaningful outcomes beyond mortality alone. In this study, Collins et al derived the "STRATFY" AHFS risk assessment tool from a final cohort of 1,033 ED patients with AHFS (Figure 2). The tool includes variables such as age, vital signs, use of oxygen/ACEIs/dialysis, ECG data, and laboratory data to assess risk.

Figure 2. Comparison of the 3 Risk Stratification Tools.[33]

EMBARGOED

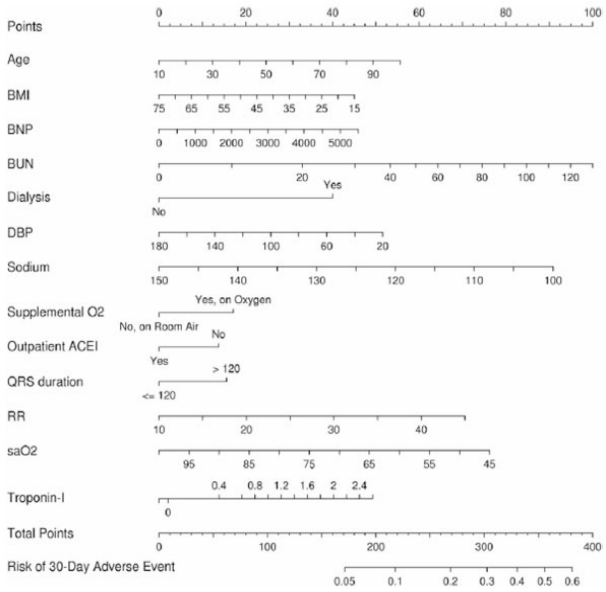


FIGURE 3. Nomogram for Use of the STRATIFY Decision Tool
Points for each variable are calculated on the upper "**Points**" bar and summed across the variables to give the total points. The total points are then found on the "Total Points" bar at the bottom, and risk is determined by drawing a perpendicular line to the "Risk of 30-Day Adverse Event" line. ACEI = angiotensin-converting enzyme inhibitor; BMI = body mass index; BNP = B-type natriuretic peptide; BUN = blood urea nitrogen; DBP = diastolic blood pressure; RR = respiratory rate; saO2 = arterial oxygen saturation.

(PENDING COPYRIGHT PERMISSION)

The disposition decision of the original cohort of patients was at the discretion of the treating physicians and occurred independently of the STRATIFY tool results. Of the total 1,033 patients, 953 (92%) were admitted to the hospital and 80 (8%) were discharged home. Patients were followed for 30 days and screened for the following adverse event outcomes: acute coronary syndrome (ACS)/percutaneous coronary intervention (PCI)/coronary artery bypass grafting (CABG), emergency dialysis, intubation, mechanical cardiac support, or death. The derived STRATIFY decision tool was used to identify patients at <1%, <3%, and <5% risk of an adverse outcome. No patients were found to be at <1% risk for an adverse outcome. However, 1.4% (N=14) were found to have a <3% risk of an adverse outcome, and 13.0% (N=134) were found to have a <5% risk of an adverse outcome. Among the 134 patients at <5% risk of an adverse event, it was determined that there was 1 death that occurred more than 5 days after the initial ED evaluation. When compared to the actual disposition of patients by the ED physicians, the authors determined that the use of the STRATIFY tool for disposition decision-making at a cutoff of <5% risk of an adverse event would have allowed for an additional 105 patients (10%) to be discharged home. This study had multiple limitations including the potential for recruitment bias resulting from the convenience sample; 63 patients withdrew from the study, 18 patients were lost to follow-up, and the tool was only internally validated using bootstrap methods so no external validation was performed. Additionally, the

EMBARGOED

application of the STRATIFY tool is extremely challenging as it requires the drawing of perpendicular lines on a nomogram to both assess the value of points for each variable as well as the overall 30-day risk of an adverse event. Finally, not only is a 5% risk of the serious adverse outcomes possibly too high to allow to occur, but the 95% upper limit of the confidence interval extends to 10%.

In the last class III paper, Lee et al[34] derive and then prospectively validate the Emergency Heart Failure Mortality Risk Grade for Acute Heart Failure (EHMRG7) for its ability to predict 7-day and 30-day mortality among ED patients with AHFS. They then compare these results with clinicians' general gestalt of mortality risk. The EHMRG7 tool was derived and first reported in an earlier paper by Lee (Figure 3).[34]

Figure 3. Comparison of the 3 Risk Stratification Tools.[34]

Variable	Units	Additive or Multiplicative Component
Age	y	2 × age
Transported by EMS	If "yes"	+60
SBP	mm Hg*	−1 × SBP
Heart rate	beats/min†	1 × heart rate
Oxygen saturation	%‡	−2 × oxygen saturation
Creatinine	mg/dL§	20 × creatinine
Potassium	4.0 to 4.5 mmol/L	0
	≥4.6 mmol/L	+30
	≤3.9 mmol/L	+5
Troponin	>ULN	+60
Active cancer	If "yes"	+45
Metolazone at home	If "yes"	+60
Adjustment factor‖		+12
Total		**EHMRG score¶**

EHMRG = Emergency Heart Failure Mortality Risk Grade; EMS = emergency medical services; SBP = systolic blood pressure; ULN = upper limit of normal.
* Initial/triage SBP, maximum of 160 mm Hg.
† Initial/triage heart rate, minimum of 80 beats/min and maximum of 120 beats/min.
‡ Lowest initial/triage oxygen saturation, maximum of 92%.
§ If creatinine concentration is in μmol/L, divide by 88.4 to convert to mg/dL.
‖ Adjustment factor of +12 added to allow for an approximate 0 median score.
¶ All variables are required to calculate the score; users are cautioned against estimating component values. The EHMRG is not for use in patients who are dialysis-dependent.

(PENDING COPYRIGHT PERMISSION)

The tool incorporates variables for age, vital signs, whether the patient was transported by emergency medical service, use of metolazone at home, the presence of cancer, and laboratory data to assess risk. Each variable has a numeric value that is further adjusted by formulae and summed to result in a final numeric score. A range of scores is then used to define 5 distinct categories of mortality risk. The tool has notably only been assessed for its ability to predict mortality risk. Both the 7- and 30-day respective mortality rates are reported for each of the 5 categories of risk are as follows: Category 1: 0.0% / 0.0%, Category 2: 0.0% / 1.9%, Category 3: 0.6% / 3.9%, Category 4: 1.9% / 5.9%, and Category 5: 3.9% / 14.3%. The study did show that physician gestalt generally overestimates the mortality risk of lower-risk patients and underestimates the mortality risk of the

EMBARGOED

highest-risk patients. For this reason, clinicians may use the tool to help prevent the disposition to home of the higher risk patients for death. However, patients assessed to be lower risk may still not be confidently discharged home since the tool did not evaluate other important outcomes beyond mortality such as acute myocardial infarction, need for cardiac intervention, need for endotracheal intubation/mechanical ventilation, need for hemodialysis, and readmission. Additionally, this study was limited by the fact that the tool was only validated internally. Given the unintuitive calculations that must be further performed for several of the tool's variables, it would seem that it would be challenging to attempt to calculate a score using this tool in a busy ED without being connected to an online EHMRG7 calculator.

Summary

To date, no study has derived an AHFS risk tool that has been used to prospectively determine an ED patient's disposition, had researchers disposition patients based solely on the results of the tool, and then followed patients over time for the development of meaningful favorable or unfavorable outcomes. Given the diverse and complex nature of ED patients with AHFS, there may never be a tool that is sufficiently accurate in assessing the risk that it may be used alone. Instead, we may more likely have tools that perform relatively well and then clinicians must employ shared decision-making strategies to determine what is best for any one patient. Still, questions remain as to what outcomes should be considered "meaningful" and what incidence of those outcomes is too high. One would think that the risk tolerance for death may be far less than the risk tolerance for readmission. Should all of these outcomes of different significance be grouped together and reported on as one event when any of them occur, or should each outcome also be reported separately with their own statistics?

Future Research

Future research should focus on developing an AHFS risk stratification tool that successfully predicts clinically important outcomes, and may be easily applied in a prospective, systematic fashion to all ED patients presenting with AHFS. Study patients would then be admitted or discharged solely on the basis of either the tool's risk stratification alone or a reproducible process that incorporates the tool's results and followed over time for clinically important outcomes. Additional studies that further assess which outcomes are most meaningful, what

EMBARGOED

746 incidence of these outcomes should be considered unacceptable, and how much admission to a hospital actually
747 prevents adverse outcomes from occurring or being associated with greater morbidity or mortality would also be
748 of value. Finally, the incorporation of prospectively validated risk modeling into formal machine learning
749 algorithms that provide clinical decision support within existing workflows may not only prove more accurate but
750 also more efficient.[35]
751

EMBARGOED

REFERENCES

1. Benjamin EJ, Muntner P, Alonso A, et al. Heart disease and stroke statistics-2019 update: a report from the American Heart Association. *Circulation*. 2019;139:e56-e528.

2. Edoute Y, Roguin A, Behar D, et al. Prospective evaluation of pulmonary edema. *Crit Care Med*. 2000;28:330-335.

3. Silvers SM, Howell JM, Kosowsky JM, et al. Clinical policy: critical issues in the evaluation and management of adult patients presenting to the emergency department with acute heart failure syndromes. *Ann Emerg Med*. 2007;49:627-669.

4. Collins S, Storrow AB, Albert NM, et al. Early management of patients with acute heart failure: state of the art and future directions. A consensus document from the society for academic emergency medicine/heart failure society of America acute heart failure working group. *J Card Fail*. 2015;21:27-43.

5. Bedetti G, Gargani L, Corbisiero A, et al. Evaluation of ultrasound lung comets by hand-held echocardiography. *Cardiovasc Ultrasound*. 2006;4:34.

6. Martindale JL, Noble VE, Liteplo A. Diagnosing pulmonary edema: lung ultrasound versus chest radiography. *Eur J Emerg Med*. 2013;20:356-360.

7. Laursen CB, Sloth E, Lassen AT, et al. Point-of-care ultrasonography in patients admitted with respiratory symptoms: a single-blind, randomised controlled trial. *Lancet Respir Med*. 2014;2:638-646.

8. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: the task force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC). Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur J Heart Fail*. 2016;18:891-975.

9. McGivery K, Atkinson P, Lewis D, 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;20:343-352.

10. Martindale JL, Wakai A, Collins SP, et al. Diagnosing acute heart failure in the emergency department: a systematic review and meta-analysis. *Acad Emerg Med*. 2016;23:223-242.

11. Staub LJ, Mazzali Biscaro RR, Kaszubowski E, et al. Lung ultrasound for the emergency diagnosis of pneumonia, acute heart failure, and exacerbations of chronic obstructive pulmonary disease/asthma in adults: a systematic review and meta-analysis. *J Emerg Med*. 2019;56:53-69.

12. Lian R, Zhang GC, Yan ST, et al. Role of ultrasound lung comets in the diagnosis of acute heart failure in emergency department: a systematic review and meta-analysis. *Biomed Environ Sci*. 2018;31:596-607.

13. Pivetta E, Goffi A, Lupia E, et al. Lung ultrasound-implemented diagnosis of acute decompensated heart failure in the ED: a SIMEU multicenter study. *Ches*t. 2015;148:202-210.

14. Pivetta E, Goffi A, Nazerian P, et al. Lung ultrasound integrated with clinical assessment for the diagnosis of acute decompensated heart failure in the emergency department: a randomized controlled trial. *Eur J Heart Fail*. 2019;21:754-766.

EMBARGOED

15. Sartini S, Frizzi J, Borselli M, et al. Which method is best for an early accurate diagnosis of acute heart failure? Comparison between lung ultrasound, chest X-ray and NT pro-BNP performance: a prospective study. *Intern Emerg Med.* 2017;12:861-869.

16. Nakornchai T, Wonkkanthai P, Monsomboon A, et al. Diagnostic accuracy of the multi-organ ultrasound performed by the novice emergency medicine residents for diagnosis of acute heart failure in Thailand. *J Med Assn Thai*. 2019;102:582-587.

17. Buessler A, Chouihed T, Duarte K, et al. Accuracy of several lung ultrasound methods for the diagnosis of acute heart failure in the ED: a multicenter prospective study. *Chest*. 2020;157:99-110.

18. Anderson JL, Adams CD, Antman EM, et al. 2011 ACCF/AHA focused update incorporated into the ACC/AHA 2007 Guidelines for the Management of Patients With Unstable Angina/Non-ST-Elevation Myocardial Infarction: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. *Circulation*. 2011;123:e426-e579.

19. Anderson KL, Jenq KY, Fields JM, et al. Diagnosing heart failure among acutely dyspneic patients with cardiac, inferior vena cava, and lung ultrasonography. *Am J Emerg Med*. 2013;31:1208-1214.

20. Wong YW, Fonarow GC, Mi X, et al. Early intravenous heart failure therapy and outcomes among older patients hospitalized for acute decompensated heart failure: findings from the Acute Decompensated Heart Failure Registry Emergency Module (ADHERE-EM). *Am Heart J*. 2013;166:349-356.

21. Levy P, Compton S, Welch R, et al. Treatment of severe decompensated heart failure with high-dose intravenous nitroglycerin: a feasibility and outcome analysis. *Ann Emerg Med*. 2007;50:144-152.

22. Collins SP, Levy PD, Martindale JL, et al. Clinical and research considerations for patients with hypertensive acute heart failure: a consensus statement from the Society for Academic Emergency Medicine and the Heart Failure Society of America Acute Heart Failure Working Group. *Acad Emerg Med*. 2016;23:922-931.

23. Mebazaa A, Yilmaz MB, Levy P, et al. Recommendations on pre-hospital & early hospital management of acute heart failure: a consensus paper from the Heart Failure Association of the European Society of Cardiology, the European Society of Emergency Medicine and the Society of Academic Emergency Medicine. *Eur J Heart Fail*. 2015;17:544-558.

24. Hunt SA, Baker DW, Chin MH, et al. ACC/AHA guidelines for the evaluation and management of chronic heart failure in the adult: executive summary a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines (Committee to Revise the 1995 Guidelines for the Evaluation and Management of Heart Failure): Developed in Collaboration With the International Society for Heart and Lung Transplantation; Endorsed by the Heart Failure Society of America. *Circulation*. 2001;104:2996-3007.

25. The Rand Corporation. *The evolving role of emergency departments in the United States*. Research Report 280, 2013. Available at: https://www.rand.org/pubs/research_reports/RR280.html. Accessed May 19, 2021.

26. Storrow AB, Jenkins CA, Self WH, et al. The burden of acute heart failure on U.S. emergency departments. *JACC Heart Fail*. 2014;2:269-277.

EMBARGOED

27. Jackson SL, Tong X, King RJ, et al. National Burden of Heart Failure Events in the United States, 2006 to 2014. *Circ Heart Fail*. 2018;11:e004873.

28. Stiell IG, McMurtry MS, McRae A, et al. The Canadian Cardiovascular Society 2018 guideline update for atrial fibrillation - A different perspective. *CJEM*. 2019;21:572-575.

29. Virani SS, Alonso A, Aparicio HJ, et al. Heart disease and stroke statistics-2021 update: a report from the American Heart Association. *Circulation*. 2021;143:e254-e743.

30. Collins SP, Pang PS. ACUTE heart failure risk stratification. *Circulation*. 2019;139:1157-1161.

31. Stiell IG, Perry JJ, Clement CM, et al. Prospective and explicit clinical validation of the Ottawa Heart Failure Risk Scale, with and without use of quantitative NT-proBNP. *Acad Emerg Med*. 2017;24:316-327.

32. Stiell IG, Clement CM, Brison RJ, et al. A risk scoring system to identify emergency department patients with heart failure at high risk for serious adverse events. *Acad Emerg Med*. 2013;20:17-26.

33. Collins SP, Jenkins CA, Harrell FE Jr, et al. Identification of emergency department patients with acute heart failure at low risk for 30-day adverse events: the STRATIFY Decision Tool. *JACC Heart Fail*. 2015;3:737-747.

34. Lee DS, Stitt A, Austin PC, et al. Prediction of heart failure mortality in emergent care: a cohort study. *Ann Intern Med*. 2012;156:767-775, w-261, w-262.

35. Sax DR, Marg DG, Huang J, et al. Use of machine learning to develop a risk-stratification tool for emergency department patients with acute heart failure. *Ann Emerg Med*. 2021;77:237-248.

36. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi: 10.1136/bmj.n71

EMBARGOED

Appendix A. Literature classification schema.*

Design/ Class	Therapy†	Diagnosis‡	Prognosis§
1	Randomized, controlled trial or meta-analysis of randomized trials	Prospective cohort using a criterion standard or meta-analysis of prospective studies	Population prospective cohort or meta-analysis of prospective studies
2	Nonrandomized trial	Retrospective observational	Retrospective cohort Case control
3	Case series	Case series	Case series

*Some designs (eg, surveys) will not fit this schema and should be assessed individually.
†Objective is to measure therapeutic efficacy comparing interventions.
‡Objective is to determine the sensitivity and specificity of diagnostic tests.
§Objective is to predict outcome, including mortality and morbidity.

Appendix B. Approach to downgrading strength of evidence.

	Design/Class		
Downgrading	**1**	**2**	**3**
None	I	II	III
1 level	II	III	X
2 levels	III	X	X
Fatally flawed	X	X	X

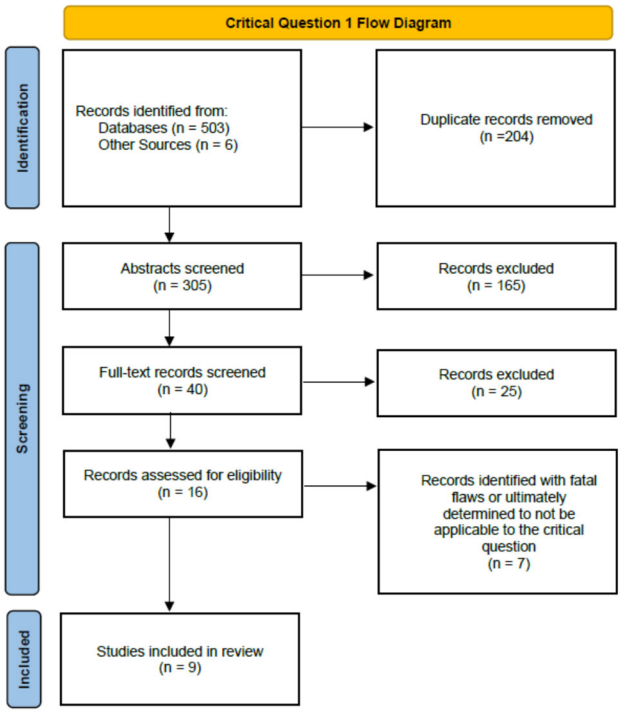
Appendix C. Likelihood ratios and number needed to treat.*

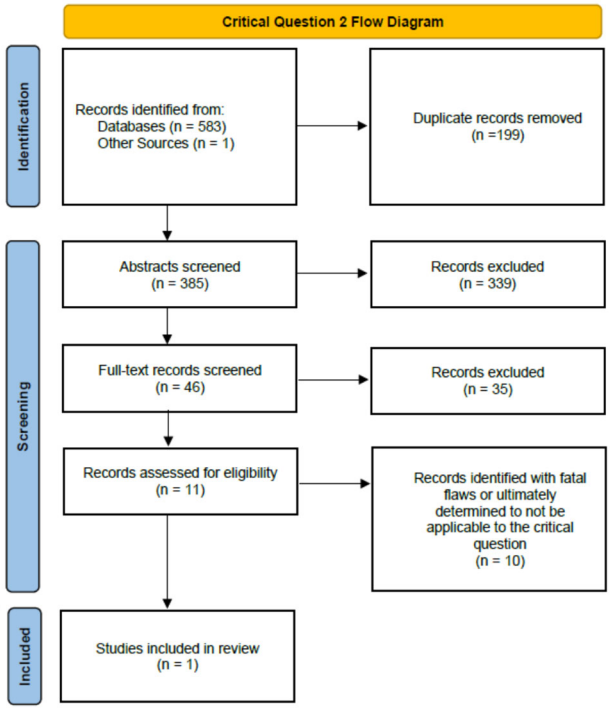
LR (+)	LR (–)	
1.0	1.0	Does not change pretest probability
1–5	0.5–1	Minimally changes pretest probability
10	0.1	May be diagnostic if the result is concordant with pretest probability
20	0.05	Usually diagnostic
100	0.01	Almost always diagnostic even in the setting of low or high pretest probability

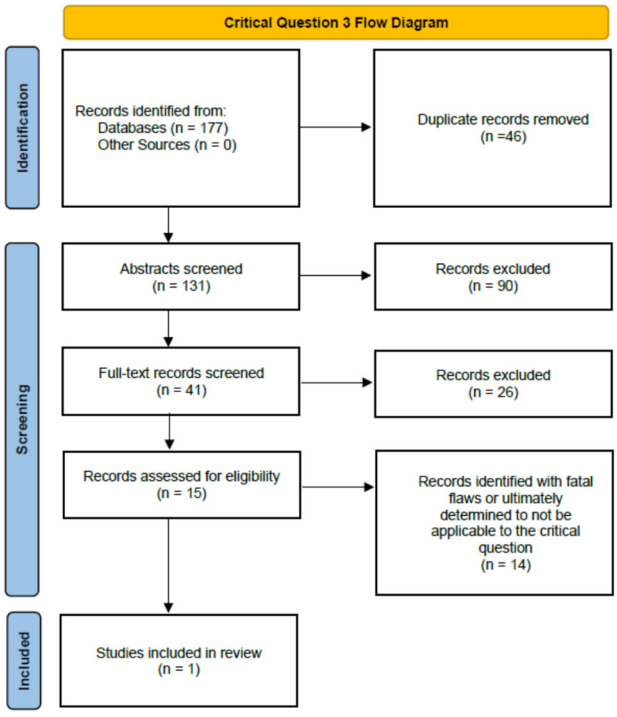
LR, likelihood ratio.
*Number needed to treat (NNT): number of patients who need to be treated to achieve 1 additional good outcome; NNT=1/absolute risk reduction×100, where absolute risk reduction is the risk difference between 2 event rates (ie, experimental and control groups).

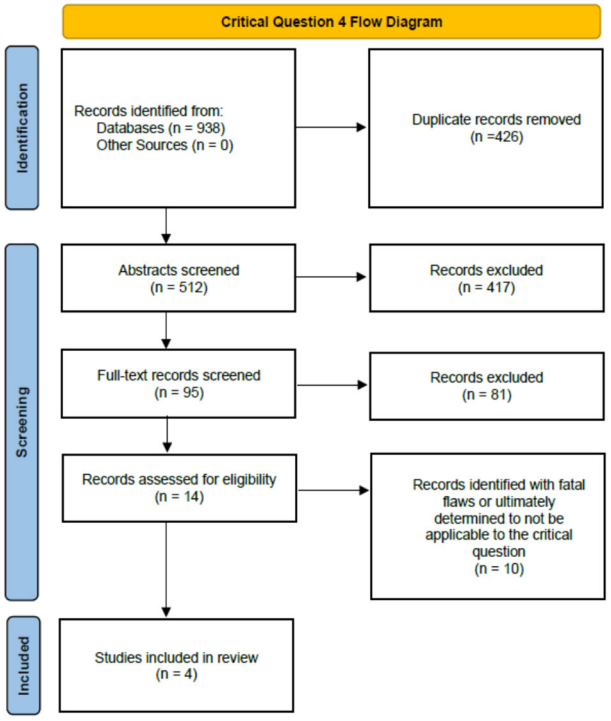
EMBARGOED

Appendix D. PRISMA[36] flow diagrams.









Evidentiary Table.

Study & Year Published	Class of Evidence	Setting & Study Design	Methods & Outcome Measures	Results	Limitations & Comments
McGivery et al[9] (2018)	II for Q1	Systematic review and meta-analysis	Prospective studies that reported on the sensitivity and specificity of B-lines in dyspneic ED patients; all included studies used at least one of the following alternate tests in their clinical diagnosis: CXR, BNP, NT-pro-BNP, or echocardiography; importantly, data from these tests were blinded from the sonographers; the bedside ultrasound was performed by emergency physicians, emergency medicine residents, ultrasound fellows, medical students, and cardiologists; all studies meeting the inclusion criteria also met the requirements for methodological quality using the CASP questionnaire	3,674 articles identified with 7 ultimately included; N=1,861; the random effects pooled results for sensitivity and specificity for ED-performed bedside LUS for the diagnosis of ADHF were 82.5% (95% CI 66.4% to 91.8%) and 83.6% (95% CI 72.4% to 90.8%), respectively; positive LR was 4.8 (95% CI 2.6 to 9.1), negative LR was 0.19 (95% CI 0.09 to 0.39)	Significant heterogeneity among included studies with large variation in study sample size
Martindale et al[10] (2016)	III for Q1	Systematic review and meta-analysis of both prospective and retrospective studies	Included both prospective and retrospective observational studies; structured searches using MeSH; used QUADAS-2 to assess quality; outcome: AHF	9,405 articles identified with 57 ultimately included; N=17,893; significant study heterogeneity, including prevalence of AHF; LUS showed pooled positive LR for AHF of 7.4	Authors pooled results although significant heterogeneity

EMERGED

Evidentiary Table. (continued)

Staub et al[11] (2019)	III for Q1	Systematic review and meta-analysis	Included both prospective and retrospective observational studies; structured searches using MeSH; used QUADAS-2 to assess quality; outcome: AHF	11,017 articles identified with 14 ultimately included related to AHF; N=2,778; AUROC 91%; sensitivities ranged from 73% to 93%, specificities ranged from 84% to 93%; positive LR ranged from 4.8 to 14, negative LR ranged from 0.07 to 0.54	Significant study heterogeneity
Lian et al[12] (2018)	III for Q1	Systematic review and meta-analysis	Included both prospective and retrospective observational studies; structured searches using MeSH; used QUADAS-2 to assess quality; outcome: AHF	8,000 articles identified in search with 15 ultimately included; N=3,309 patients; significant heterogeneity among findings; pooled sensitivity 85%; range across studies: 33% to 100%; pooled specificity 91%; range across studies: 54% to 100%; positive LR 8.9, negative LR 0.14	Most included studies were deemed low risk of bias, although details of this assessment are sparse; authors pooled results although significant heterogeneity
Pivetta et al[14] (2019)	III for Q1	Randomized clinical trial; two emergency departments	Included adult ED patients 18 years of age or older with acute dyspnea, stratified by presumptive etiology (AHF or non-AHF); participants were then randomized to either LUS or CXR/NT-proBNP; outcome: AHF as independently assessed by 2 physicians blinded to allocation	N=518; AUROC for LUS 0.95, AUROC for CXR/NT-proBNP 0.87, AUROC for clinical evaluation along 0.85	Limited generalizability due to 2 sites and LUS performed by specified study emergency physicians

EMBARGOED

Evidentiary Table. (continued)

Pivetta et al[13] (2015)	III for Q1	Prospective observational study; multiple centers in Italy	Included adult patients 18 years of age or older who presented to the ED with acute dyspnea; EPs assessed clinical diagnosis of AHF and performed LUS; outcomes: AHF as determined by independent review of medical record by 2 physicians blinded to US results	N=1,005; LUS sensitivity 97% (95% CI 95% to 98%), LUS specificity 97% (95% CI 96% to 99%]); initial clinical workup without US sensitivity 85% (95% CI 82% to 88%) and specificity, 90% (95% CI 87% to 92%); CXR alone sensitivity 70% (95% CI 65% to 74%) and specificity 82% (95% CI 79% to 85%)	Large multi-center cohort; possible selection bias; emergency physicians had specific training, so possibly not generalizable to broad emergency care practice
Sartini et al[15] (2017)	III for Q1	Prospective observational study; 1 hospital	Included adult ED patients 18 years of age or older presenting with acute dyspnea not related to trauma; LUS performed by dedicated study-specific emergency physicians; outcome: AHF as determined by an independent panel of experts, including cardiology and emergency medicine	N=236; 48% with AHF <u>LUS</u> Sensitivity 58% Specificity 88% <u>CXR</u> Sensitivity 75% Specificity 86% <u>NT-proBNP</u> Sensitivity 96% Specificity 28%	Limited generalizability due to small sample size and single institution; possible spectrum bias
Nakornchai et al[16] (2019)	III for Q1	Prospective observational study; single center, large urban, tertiary care center in Thailand	Included adult patients 18 year of age or older with acute dyspnea and with AHF as part of the differential; EM resident blinded to patient information performed US; outcome: AHF as determined by 2 emergency physicians blinded to the US results	N=62; 65% were diagnosed with AHF; sensitivity 60%, specificity 73%	Small sample size; limited generalizability; possible selection bias

EMBARGOED

Evidentiary Table. (continued)

Buessler et al[17] (2020)	III for Q1	Prospective, multi-center observational study	Patients >50 years of age who were admitted for acute dyspnea and for whom the physician had diagnostic uncertainty; excluded patients who experienced trauma or who had systolic BP <70 mm Hg; 4-point, 6-, 8-, and 28-point LUS was performed by ultrasound-certified emergency physicians, as well as clinical assessment using the BREST score; outcomes: final AHF diagnosis at discharge, adjudicated by 2 physicians and blinded to US results	N=117; N=69 with AHF; among the 69 patients the 4-, 6-, 8-, and 28-point LUS identified AHF in 27%, 56%, 55%, and 77% of patients, respectively; C-index was: 73% for the Brest score; 64% for 4-point; 72% for 6-point; 74% for 8-point, and; 72% for 28-point, individually C-index for each increased from 3.5 to 7.3 when added to BREST score with p-values ranging from 0.1 to 0.004.	Potential selection bias; US performed by trained, certified, physicians, which may limit generalizability; small sample although heterogeneous clinical sites
Wong et al [20] (2013)	III for Q2	Secondary analysis of the ADHERE-EM registry, US Centers for Medicare and Medicaid, 2004 to 2005, across 83 hospitals	Patients 65 years of age or older who were hospitalized with a primary or secondary diagnosis of HF; Cox proportional hazard model to assess association of time to treatment with a composite 30-day all-cause mortality or readmission	N=6,971; median time to first IV therapy was 2.3 hours (interquartile range of 1.1 to 4.4); 30-day all-cause mortality or readmission was 27.4%; time to treatment was not associated with increased risk of composite 30-day mortality or re-admission (hazard ratio 1.00 [95% CI 1.0 to 1.0]); every hour delay in treatment was associated with risk of in hospital mortality (adjusted odds ratio 1.01 [95% CI 1.00 to 1.02])	Secondary analysis of an existing dataset; multi-center; selection bias possible given inclusion of only older patients

EMBARGOED

Evidentiary Table. (continued)

Levy et al[21] (2007)	III for Q3	Prospective quasi-experiment, described as a nonrandomized, open-label study; two institutions in Detroit	Adult patients 18 years of age or older with a SBP >160 mm Hg or mean arterial blood pressure >120 mm Hg; exclusions included noncardiac pulmonary edema, need for immediate intubation, or CPR; initial treatment included nonrebreather mask oxygen, sublingual NTG and furosemide; patients deemed to require IV NTG were approached for inclusion; HD IV NTG included 0.3 to 0.5 mcg/kg/min and titrated to maximum of 400 mcg/min, but with initial bolus of 2,000 mcg with subsequent 2,000 mcg boluses every 3 to 5 min at discretion of treating emergency physician; effectiveness outcome: Intubation within 6 hours of treatment; safety outcomes: neurological or cardiovascular complications; secondary outcomes: BiPAP, need for ICU, hospital length of stay, renal dysfunction, 30-day ED recidivism	N=64; N=29 HD NTG; Mean dose of IV NTG 6,500 mcg <u>Intubation</u> HD NTG 14% Non-HD NTG 27% <u>BiPAP</u> HD NTG 7% Non-HD NTG 20% <u>ICU Admission</u> HD NTG 38% Non-HD NTG 80% <u>Symptomatic Hypotension</u> HD NTG 3% Non-HD NTG 0% <u>Cardiac Ischemia by Biomarker</u> HD NTG 17% Non-HD NTG 29%	Small sample size; significant imbalances between study groups; control group data obtained retrospectively; no adjustment for confounding

EMBARGOED

Evidentiary Table. (continued)

Stiell et al[31] (2017)	II for Q4	Multi-center, prospective cohort study; 6 Canadian EDs	Enrollment included a sample of ED patients >50 years of age presenting with dyspnea due to AHF; outcomes: SAE defined as mortality within 30 days of ED visit, or admission, intubation, acute myocardial infarction, major procedure within 14 days	N=1,100; SAEs occurred in 170 (15.5%); prognostic accuracy of the OHFRS was: OHFRS >1 91.8% sensitivity and 24.9% specificity, and when NT-BNP included, OHFRS >1, 95.8% sensitivity and 13.6% specificity for identifying SAE	Potential for intention bias related to admitted patients who may not have experienced SAE simply because they were admitted; potential for selection bias related to convenience enrollment; not all patients had NT-BNP measurements; use of same EDs that were involved in the development of the OHFRS (see Stiell et al 2013)
Stiell et al[32] (2013)	III for Q4	Multi-center, prospective cohort study; 6 Canadian EDs	Enrollment included a convenience sample of ED patients ≥50 y presenting with dyspnea due to ADHF exacerbation; outcome: SAE defined as mortality of any cause within 30 days of the ED visit, or admission, intubation, acute myocardial infarction, major procedure, or relapse within 14 days of the ED visit	N=559; SAEs occurred in 65 (11.6%) and in only 31 (5.5%) who were not admitted to the hospital; The OHFRS was developed and included 10 characteristics with SAEs ranging from 2.8% for a Score=0, and 89.0% for a Score=9 with good calibration (Hosmer-Lemeshow goodness-of-fit p=0.95) and discrimination (AUROC of 0.75)	Potential for selection bias given convenience sampling; OHFRS only internally validated using bootstrap methods; thus, no external validation performed
Collins et al[33] (2015)	III for Q4	Multi-center, prospective cohort study; 4 EDs in the United States	Enrollment included adult patients presenting to the ED with acute HF using the modified Framingham criteria; outcomes: SAE within 30 days, defined as all-cause mortality, acute coronary syndrome, CPR, mechanical cardiac support, intubation, hemodialysis, or need for percutaneous coronary intervention	N=1,033; adverse event occurred in 126 (12%); The STRATIFY decision tool had moderate discrimination (*c* statistic 0.68) and good calibration; a score of 5 resulted in a sensitivity of 95% and specificity of 14% for severe adverse event	Potential selection bias given convenience sampling; 63 participants withdrew and 18 were lost to follow-up; the STRATIFY decision tool was only internally validated using bootstrap methods; thus, no external validation was performed

EMBARGOED

Evidentiary Table. (continued)

Lee et al[34] (2019)	III for Q4	Multi-center, prospective cohort study; 9 Canadian EDs	Enrollment included adult patients presenting to the ED with AHF; outcomes: mortality at 7 days; mortality at 30 days	N=1,983; mortality: 39 (2.0%) at 7 days and 138 (7.0%) at 30 days; compared to physician estimation, Emergency Heart failure Mortality Risk Grade (EHMRG7) had improved discrimination (*c* statistic 0.81 vs 0.71)	Mortality was the only outcome; thus, other important outcomes not assessed

ADHF, acute decompensated heart failure; *AHF*, acute heart failure; *AHFS*, Acute Heart Failure Syndromes; *AUROC*, area under the receiver operating characteristics; *BiPAP*, bilevel positive airway pressure; *BNP*, B-type natriuretic peptide; *CI*, confidence interval; *CXR*, chest x-ray; *ED*, emergency department; *h*, hour; *HD*, high-dose; *IV*, intravenous; *kg*, kilogram; *LR*, likelihood ratio; *LUS*, lung ultrasound; *mcg*, microgram; *MeSH*, Medical Subject Heading; *mg*, milligram; *min*, minute; *NPV*, negative predictive value; *NTG*, nitroglycerin; *NT-proBNP*, N-terminal pro-B-type natriuretic peptide; *OHFRS*, Ottawa Heart Failure Risk Scale; *PPV*, positive predictive value; *QUADAS-2*, Quality Assessment of Diagnostic Accuracy Studies 2; *SAE*, serious adverse event; *US*, ultrasound; *y*, year.

EMBARGOED