A proposal to standardize dyspnoea measurement in clinical trials of acute heart failure syndromes: the need for a uniform approach

Peter S. Pang1, John G.F. Cleland2*, John R. Teerlink3, Sean P. Collins4, Christopher J. Lindell5, George Sopko5, W. Frank Peacock6, Gregg C. Fonarow7, Amer Z. Aldeen1, J. Douglas Kirk8, Alan B. Storrow9, Miguel Tavares10, Alexandre Mebazaa11, Edmond Roland12, Barry M. Massie3, Alan S. Maisel13, Michel Komajda14,15, Gerasimos Filippatos16, and Mihai Gheorghiade17 for the Acute Heart Failure Syndromes International Working Group

1Department of Emergency Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; 2Department of Cardiology, Castle Hill Hospital, University of Hull, Kingston-upon-Hull, UK; 3Section of Cardiology, San Francisco Veterans Affairs Medical Center and School of Medicine, University of California San Francisco, San Francisco, CA, USA; 4Department of Emergency Medicine, University of Cincinnati, Cincinnati, OH, USA; 5National Institutes of Health, Bethesda, MD, USA; 6Department of Emergency Medicine, The Cleveland Clinic, Cleveland, OH, USA; 7Division of Cardiology, Department of Medicine, David Geffen School of Medicine at UCLA, University of California Los Angeles, CA, USA; 8Department of Emergency Medicine, University of California, Davis, Medical Center, Sacramento, CA, USA; 9Department of Emergency Medicine, Vanderbilt University Medical Center, Nashville, TN, USA; 10Department of Anesthesiology and Critical Care, Hospital Geral de Santo António, Porto, Portugal; 11Department of Anesthesiology and Critical Care Medicine, Hopital Lariboisiere, Paris, France; 12Agence Francaise de Securite Sanitaire des Produits de Sante, Paris, France; 13Division of Cardiology, Department of Medicine, Veterans Affairs San Diego Health Care System, University of California San Diego, San Diego, CA, USA; 14Department of Cardiology, University Pierre Marie Curie, Paris, France; 15Pitie Salpetriere Hospital, Paris, France; 16Department of Cardiology, Heart Failure Unit, Athens University Hospital, Attikon, Greece; 17Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Keywords
Heart failure • Clinical trials • Dyspnoea

Introduction

Dyspnoea is the most common presenting symptom amongst patients with acute heart failure syndromes (AHFS). It is distressing to patients and therefore an important target for treatment in clinical practice, clinical trials, and for regulatory approval of novel agents. Despite its importance as a treatment target, no consensus exists on how to assess dyspnoea in this setting. There is considerable uncertainty about the reproducibility of the various instruments used to measure dyspnoea, their ability to reflect changes in symptoms and whether they accurately reflect the patient’s experience. Little attempt has been made to ensure consistent implementation with respect to patients’ posture during assessment or timing in relationship to therapy. There is also limited understanding of how rapidly and completely dyspnoea responds to standard therapy. A standardized method with which to assess dyspnoea is required for clinical trials of AHFS in order to ensure uniform collection of data on a key endpoint. We propose the Provocative Dyspnoea Assessment, a method of measurement that combines sequential dyspnoea provocation by positioning and walking with a dyspnoea self assessment using a five-point Likert scale, to yield a final Dyspnoea Severity Score. This proposed tool requires detailed validation but has face validity for the uniform assessment of dyspnoea.
Dyspnoea measurement in acute heart failure syndromes clinical trials

and that is also sensitive and specific to changes in dyspnoea in AHFS, does not currently exist. Furthermore, a uniform methodology or set of conditions under which dyspnoea is assessed is also absent.7 Lack of standardization may introduce unwanted variability around a key target and confound attempts to demonstrate the efficacy of new treatments for AHFS.

Although dyspnoea is an important aspect of AHFS, investigators are often distracted by technical aspects of the studies they conduct and may pay less attention to something as apparently simple as measuring dyspnoea. Ensuring investigators take time and care over the measurement of this primary outcome measure in a study seems a fundamentally good idea. Accordingly, an international group of cardiologists, intensive- and emergency-care physicians, hospital physicians, representatives from regulatory agencies, and the National Institutes of Health of the United States of America have proposed that a uniform method of dyspnoea assessment is needed, and they have developed an initial consensus proposal for dyspnoea measurement in clinical trials of AHFS.5

Definition of dyspnoea

The American Thoracic Society defines dyspnoea as ‘a subjective experience of breathing discomfort that consists of qualitatively distinct sensations that vary in intensity. The experience derives from interactions among multiple physiological, psychological, social, and environmental factors, and may induce secondary physiological and behavioral responses.8 The European Society of Cardiology (ESC) guidelines on acute heart failure do not provide a definition of dyspnoea, presumably assuming that clinicians know what it means.9 Braunwald simply defines dyspnoea as an abnormally uncomfortable awareness of breathing.10

The fact that a precise understanding of the symptom of dyspnoea from a pathophysiologic or psychosocial perspective has not been fully elucidated compounds the difficulty of dyspnoea measurement.6 Other complicating factors include the difference in a patient’s subjective experience of dyspnoea depending on the underlying disease (i.e. obstructive lung disease vs. heart failure (HF)).11,12 In addition, the causes of dyspnoea induced by exertion in patients with chronic HF may differ from those causing acute dyspnoea in patients with AHFS.

Dyspnoea is the patient’s subjective experience of disease for which no objective measure is an adequate substitute. Exploration of the subjective nature and severity of distress that dyspnoea causes, standardization of the conditions under which questions are asked, knowledge of the underlying disease causing the symptom, and examination of the context (de novo, acute-on-chronic) must be considered. Moreover, the aetiology, physiological state (i.e. posture, rest, exertion, etc.) degree and speed of change, duration, psychological status, and patient experience and expectations might help ensure a more robust, although still subjective assessment of dyspnoea.

Dyspnoea as a manifestation of pulmonary congestion in acute heart failure syndromes

In AHFS, dyspnoea most likely reflects pulmonary congestion due to an increase in pulmonary capillary wedge pressure (PCWP) as a result of left ventricular (LV) dysfunction (systolic and/or diastolic) and/or valvular abnormalities.13 Increases in PCWP indicating haemodynamic congestion, by itself, may or may not cause dyspnoea.13 Patients with chronically elevated PCWP may not have severe symptoms or radiographic pulmonary congestion, while a patient with new onset HF and similar PCWP may have severe symptoms and overt radiographic pulmonary oedema.14,15 The trigger point at which haemodynamic congestion becomes clinically overt depends on the extent and rate at which PCWP rises, oncocytic pressure, and compensatory mechanisms such as the efficiency of pulmonary lymphatic drainage, the permeability of the alveolar-capillary membrane, and the presence or absence of underlying lung disease that reduces pulmonary reserve.13 Therefore, the PCWP that induces symptoms will vary amongst patients but be more constant, at least over a period of days, within an individual.

Epidemiology of dyspnoea in large surveys of acute heart failure syndromes

Dyspnoea is the most common reason why patients with AHFS present for medical care (Table 1). In the Acute Decompensated Heart Failure National Registry (ADHERE) and Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF), registries approximately 90% of patients reported dyspnoea at rest or on exertion. However, only 44% of patients in OPTIMIZE-HF reported dyspnoea at rest, which is similar to the proportion reporting dyspnoea at rest at the time of admission in the EuroHeart Failure Survey-I (EHFS-I), a survey of deaths and discharges rather than of admissions. Despite the prominence of dyspnoea, the clinical presentation of overt pulmonary oedema is present in a minority of patients. In the EuroHeart Failure Survey-II (EHFS-II), where dyspnoea due to AHFS was an entry criterion, only 16% were reported to have pulmonary oedema, as defined by alveolar oedema by chest X-ray or with oxygen saturation < 90% (without supplemental oxygen).

Regulatory perspectives on dyspnoea as an endpoint

The perception of AHFS has changed substantially over the last 5 years, with AHFS now recognized as a distinct entity within the pathophysiological spectrum of HF.5 The steps required for regulatory approval have also evolved.1 When milrinone was approved by the FDA in 1987, the demonstration of haemodynamic improvement was considered sufficient evidence of efficacy.4 Nesiritide was approved in 2002 after demonstrating favourable haemodynamics, safety, and symptom (i.e. dyspnoea) improvement, even for a very short period. However, post-approval questions regarding the safety of nesiritide19–21 have raised the bar considerably with regard to safety.

Regulators in the USA and Europe concur that, as for chronic heart failure, endpoints in clinical trials must be relevant to patient welfare and might include symptoms, length of hospitalization, re-admission, disabling morbid events, or death.

Regulators request that the instrument for assessing dyspnoea should ideally be well validated, clinically justified, and defined a priori. A global assessment of the patient’s clinical status may
be useful, complementing clinical information and as a co-primary endpoint. Unfortunately, regulators provide no specific guidance on how dyspnoea should be assessed. No validated ‘gold-standard’ instrument exists. The lack of investment in developing a methodology for measuring dyspnoea has led to the use of an assortment of unvalidated, poorly standardized scales (Table 2) that have occasionally demonstrated statistical differences between interventions that are of questionable clinical relevance and possibly biased by the results of haemodynamic monitoring. Whether the scales used, so far, have been accurate and the treatments ineffective, or the treatments effective but the tools of assessment inadequate is uncertain. Accordingly, investigators are free to use any method of assessing dyspnoea that is clinically justified and prospectively defined, since none are validated.

Novel treatment will likely be required to demonstrate some benefit over standard treatment for AHFS, including oxygen, diuretics, morphine, and vasodilators. Dyspnoea usually responds rapidly to these treatments, leaving a narrow window of opportunity to demonstrate an effect from new drugs. Whether the scales used, so far, have been accurate and the treatments ineffective, or the treatments effective but the tools of assessment inadequate is uncertain. Accordingly, investigators are free to use any method of assessing dyspnoea that is clinically justified and prospectively defined, since none are validated.

Timing of dyspnoea measurement from recent acute heart failure syndromes trials

Most AHFS trials measure dyspnoea after hospital admission, well after initial therapies have been administered. These same trials demonstrate dyspnoea improves rapidly and substantially with conventional therapy. However, initial conventional therapy is often poorly documented and, although perhaps consistent within individual institutions, there is no widely accepted standard therapeutic approach. The timing and dose of all therapies given
<table>
<thead>
<tr>
<th>Trial and intervention</th>
<th>Year</th>
<th>Number of patients</th>
<th>Was dyspnoea an endpoint?</th>
<th>Who measured dyspnoea?</th>
<th>When was patient enrolled?</th>
<th>When was dyspnoea measured?</th>
<th>Instrument or scale used to measure dyspnoea</th>
<th>Detailed description of how dyspnoea was measured</th>
<th>Improvement in dyspnoea</th>
</tr>
</thead>
<tbody>
<tr>
<td>VERITAS22,24,29 (tezosentan (T) vs. placebo (P))</td>
<td>2005</td>
<td>1448 (1760)*</td>
<td>Yes (primary)</td>
<td>Patient</td>
<td>After admission</td>
<td>0, 3, 6, and 24 h (AUC)</td>
<td>VAS at all three timepoints (AUC)</td>
<td>Not specified</td>
<td>No difference between T and P in two parallel studies.</td>
</tr>
<tr>
<td>RITZ-125,26 (tezosentan (T) vs. placebo (P))</td>
<td>2001</td>
<td>669</td>
<td>Yes (primary)</td>
<td>Patient</td>
<td>After admission</td>
<td>0, 3, 6, and 24 h</td>
<td>7 point Likert scale</td>
<td>NA</td>
<td>No difference between tezosentan vs. placebo</td>
</tr>
<tr>
<td>RITZ-236 (tezosentan (T) vs. placebo (P))</td>
<td>2003</td>
<td>292</td>
<td>Yes (co-primary with cardiac index)</td>
<td>Patient</td>
<td>After admission</td>
<td>0, 6, and 24 h</td>
<td>7 point Likert scale</td>
<td>Not specified</td>
<td>6 h, T vs. P (P=0.46); 24 h, T (50 mg/h) 44%, T (100 mg/h) 49% vs. P 31% (P=0.048)</td>
</tr>
<tr>
<td>OPTIME-CHF37 (milrinone (M) vs. placebo (P))</td>
<td>2002</td>
<td>951</td>
<td>Yesb (secondary)</td>
<td>Patient</td>
<td>After admission</td>
<td>0, 3 days, discharge</td>
<td>Dyspnoea was part of a non-validated composite score</td>
<td>Not specified</td>
<td>No difference observed</td>
</tr>
<tr>
<td>VMAC23 (nesiritide (NES) vs. nitrates (NTG) vs. placebo (P))</td>
<td>2002</td>
<td>489</td>
<td>Yes (co-primary with PCWP)</td>
<td>Patient</td>
<td>After admission</td>
<td>0, 3, 6, and 24 h</td>
<td>7 point Likert scale</td>
<td>Not specified</td>
<td>3 h, NES 76 vs. P 62% (P=0.03); NTG 73 vs. P 62% (P=0.19); NES vs. NTG (P=0.56); 24 h, NES 90 vs. NTG 85% (P=0.13)</td>
</tr>
<tr>
<td>EVEREST22 (tolvaptan (T) vs. placebo (P))</td>
<td>2007</td>
<td>Trial A=2048, Trial B=2085</td>
<td>Yes (secondary)</td>
<td>Patient</td>
<td>After admission</td>
<td>Inpatient day 1</td>
<td>7 point scale</td>
<td>Not specified</td>
<td>Trial A: T 77 vs. P 71% (P&lt;.001); Trial B: T 72 vs. P 65% (P&lt;.001)</td>
</tr>
<tr>
<td>SURVIVE24 (levosimendan (L) vs. dobutamine (D))</td>
<td>2007</td>
<td>1327</td>
<td>Yes (secondary)</td>
<td>Patient</td>
<td>After admission</td>
<td>24 h</td>
<td>7 point Scale</td>
<td>Not specified</td>
<td>No difference between levosimendan vs. dobutamine</td>
</tr>
<tr>
<td>REVIVE-II25 (levosimendan (L) vs. placebo (P))</td>
<td>2005</td>
<td>600</td>
<td>Yes (secondary)</td>
<td>Patient</td>
<td>After admission</td>
<td>6 h</td>
<td>7 point Scale</td>
<td>NA</td>
<td>L ~66 vs. P ~55% (P=0.078) (REVIVE-II only)—estimate from abstract graph</td>
</tr>
</tbody>
</table>

VAS, visual analog scale; NA, not available (information generated from abstract only).

*VERITAS was discontinued before full enrolment due to improbability of achieving significant treatment effect. The n listed is the planned enrolment.
for measuring dyspnoea in settings other than AHFS.4,7 Ideally, use of Likert or visual analogue scales (VAS) have been validated as optimal instruments with which dyspnoea should be measured. To date, AHFS studies have not convincingly demonstrated the which dyspnoea measurement scale should be used to measure dyspnoea? (i) when should dyspnoea be measured and subsequently re-assessed? (ii) under what conditions should dyspnoea be measured? (iii) how or under what conditions should dyspnoea be measured? (iv) when should dyspnoea be measured and subsequently re-assessed?

Recommendations for the measurement of dyspnoea in future clinical trials

Assessment of dyspnoea in a uniform manner in patients with AHFS is needed. Four issues are of major importance (i) which scale should be used to measure dyspnoea? (ii) who is being assessed? (iii) how or under what conditions should dyspnoea be measured? (iv) when should dyspnoea be measured and subsequently re-assessed?

Which dyspnoea measurement scale should be used

To date, AHFS studies have not convincingly demonstrated the optimal instruments with which dyspnoea should be measured. Use of Likert or visual analogue scales (VAS) have been validated for measuring dyspnoea in settings other than AHFS.4,7 Ideally, development of a valid instrument should be the first step towards adopting a standardized method of measurement. Scales should be tested for inter- and intra-subject variability, as the patient will be the observer in all cases. The scale should be shown to be sensitive to small changes in dyspnoea. Not only should absolute values of dyspnoea (ordinal, categorical) be investigated, but also changes in the setting of dyspnoea (lying, sitting, etc). The scale should also have reasonable face validity, reflecting a clinically important change in dyspnoea status. Given the absence of a ‘gold standard’ against which to validate a dyspnoea instrument and the subjective nature of dyspnoea, psychometric evaluation may help establish a dyspnoea scale. However, despite the absence of this critical first step, trials have assessed dyspnoea and will likely continue to do so. Clarity and specificity of what is being measured and in whom, and the timing of measurement as well as the conditions under which measurement occurs is needed. We propose investigators begin to measure dyspnoea in a standardized manner to establish uniformity and maximize comparability of clinical trial data as we strive to develop a validated scale.

Dyspnoea must be cardiac in origin

For the purposes of AHFS clinical trials, dyspnoea should be secondary to heart failure so that therapy treating the underlying cause can result in improved symptoms. Requiring an elevated BNP/NT-proBNP and/or other agreed upon clinical signs and/or radiographic evidence (or haemodynamic criteria for haemodynamic studies) as inclusion criteria can facilitate patient selection by adding objective evidence that dyspnoea is due to AHFS and not another cause. BNP is appealing as it is released by the myocyte by the same stimulus that helps create the sensation of dyspnoea, that is elevated filling pressures emanating from pressure and/or volume overload.30 BNP levels are able to separate cardiac from non-cardiac causes of dyspnoea with an accuracy of 90%.31 Experience from recent trials confirms the necessity of these objective measures in addition to the more conventional, but less specific or reproducible, findings of elevated jugular venous pressure, rales, or oedema.

When should dyspnoea be measured

Expert consensus suggests that clinical trials be designed in accordance with a patients’ time phase of management.5 Stage A trials target initial presentation (i.e. the ED). Stage B trials, representing the majority of clinical trials to date, target patients with persistent signs and symptoms or who develop AHFS while in the hospital. Stage C trials focus on improving post-discharge mortality and morbidity (readmission) by optimizing therapies in hospital and/or ensuring excellent follow up care. Whilst some overlap will occur, a clear distinction between dyspnoea that is present upon arrival and dyspnoea that persists after 48 h of therapy needs to be recognized. Although dyspnoea is the most common reason for presentation, with rare exception,52 very few major trials have been conducted during Stage A of presentation.

As well as the stage of study, change in symptoms will be determined in part by when a baseline is measured and thus trials should specify the timing of intervention (e.g. before standard therapy, with standard therapy, after failure of standard therapy). As dyspnoea improves substantially with traditional therapies, only a short-time window exists in which to capture meaningful change. Thus, for Stage A trials, dyspnoea should be measured when it is most severe, which is usually in the ambulance or immediately upon presentation. Stage B trials will be designed as rescue or second-line therapy for patients who have persistent signs and/or symptoms despite initial therapy or who develop or have worsening of AHFS in the hospital setting.

Assessments of dyspnoea should be repeated but the optimal interval between assessments is uncertain. The consensus group recommended hourly assessment for the first 3 h, to identify the initial speed of response, then an assessment at 6 h and every 6 h thereafter until 24 h, to identify late response and relapses. Patients who are sleeping at the time of assessment will need to be awake for at least 10 min prior to evaluation.

Trials allow a time window for enrolment, i.e. 24–48 h after admission.7,22,23 In general, this time frame does not include the time spent from presentation (time spent in the ED) until admission. These time differences will effect the amount of therapy offered prior to actual enrolment. Further, the effect of early compared to late randomization has not been reported even though patients enrolled early have not had the presumed benefit of treatment that patients enrolled late may have had.

In addition, trials often measure dyspnoea at baseline and at fixed time points after enrolment (Table 2), which may fail to capture the fluctuating course of symptoms. The effect of some treatments, such as opiates, wear off after a few hours, leading to a recrudescence of dyspnoea, and complications such as atrial fibrillation, ischaemia, or infection can make a patient who initially responded, relapse. Measurement of symptoms at multiple times to construct an average response, using area-under-the-curve methodology is a useful approach.29 Patients who improve quickly and maintain their improvement will score highly, while slow response or relapse leads to lower scores. Issues remain about how often measurements should be made and for how long; however, most patients who are going to improve substantially will have done so within 24–48 h.

Trials allow a time window for enrolment, i.e. 24–48 h after admission.7,22,23 In general, this time frame does not include the time spent from presentation (time spent in the ED) until admission. These time differences will effect the amount of therapy offered prior to actual enrolment. Further, the effect of early compared to late randomization has not been reported even though patients enrolled early have not had the presumed benefit of treatment that patients enrolled late may have had.

In addition, trials often measure dyspnoea at baseline and at fixed time points after enrolment (Table 2), which may fail to capture the fluctuating course of symptoms. The effect of some treatments, such as opiates, wear off after a few hours, leading to a recrudescence of dyspnoea, and complications such as atrial fibrillation, ischaemia, or infection can make a patient who initially responded, relapse. Measurement of symptoms at multiple times to construct an average response, using area-under-the-curve methodology is a useful approach.29 Patients who improve quickly and maintain their improvement will score highly, while slow response or relapse leads to lower scores. Issues remain about how often measurements should be made and for how long; however, most patients who are going to improve substantially will have done so within 24–48 h.
How to measure changes in dyspnoea
As well as an absolute measurement of dyspnoea, it is important in AHFS trials to measure change. Assessing change in symptoms can be done in at least two ways. One method is to use the same instrument at baseline and repeatedly thereafter to measure symptom severity at each time-point. Change can then be inferred by computation. Another method is to ask patients whether they have changed or not. The problem with the second, seemingly simple, approach is being sure about when the patient is measuring change from, and whether the patient’s recollection is accurate. Anecdotally, many patients will report change from their most severe symptoms rather than the time of enrolment, which is hardly surprising given how ill many of these patients are. This gets even more complex when more than one follow-up measurement is made. The consensus group agreed that the former approach should be preferred, although it has been argued that patients acclimatize to being dyspnoeic, which could confound absolute measurements.

The next issue is under what conditions measurements should be made. Two patients may have moderate dyspnoea but one is sitting upright with high-flow oxygen and the other lying flat without oxygen. Few people would consider these patients equally breathless or sick. Patients with AHFS should be assessed in a standardized set of conditions ranging from sitting upright with oxygen to lying flat without oxygen, if they are able. An example of such a standardization is given in Table 3, where an ordered approach to assessing dyspnoea is suggested. Here, assessment continues across a series of conditions that are increasingly difficult for a patient to tolerate. The assessment of dyspnoea then becomes a “failure task”: which condition the patient cannot tolerate. This is similar to many psychophysical threshold assessments, which change an objectively measurable stimulus until a patient responds. The consensus group has termed this scale the Provocative Dyspnoea Assessment (PDA) scale and suggests it may provide a robust profile of dyspnoea that is sensitive to change. Further, the scale possesses face validity in that it may be likened to an adapted conventional exercise test for patients who are highly symptomatic.

Assessment of the feasibility and performance of the PDA is planned. However, since there is no valid existing tool for assessing dyspnoea in AHFS, the consensus group hopes that the PDA will become the new interim standard tool, until such time as it is adopted in a validated form or supplanted by a superior tool. Trial design and purpose will dictate how many steps of the PDA would be utilized. For example, trials that target initial presentation (Stage A) might use only the first three steps of the PDA. Trial design should also stipulate what severity of dyspnoea on this scale is required for inclusion in the study. The PDA may expand the number of patients who are eligible for studies since some patients will have moderate or severe symptoms provoked by lying flat who are nonetheless comfortable sitting upright.

Recommendations for practical application of the PDA—Dyspnoea Severity Score
As well as broad categories of failure, it is possible to refine the PDA using Likert scales or VAS to assess the severity of dyspnoea at each completed stage. Such a refinement provides a less coarse gradation of dyspnoea and might better permit assessing change. Outlined is our recommended approach to use of the PDA in an AHFS clinical trial.

Consider the PDA used in conjunction with a 5-point Likert scale, with one being the greatest severity of patients’ self-reported dyspnoea (note that a VAS or alternative Likert scale could be used). Two numbers can be reported from this study, one from the Likert, the other based on the final step on the PDA that the patient completed. These can be combined to yield a final Dyspnoea Severity Score (DSS). An example of such a combination is shown in Table 4.

This approach has neither been methodologically nor statistically validated as yet, but is proposed as a first step in such a validation process and provides uniformity of dyspnoea assessment. Also, while not yet evaluated as a scoring mechanism, several elements will benefit from standardization. First, failure of a step should be similar across studies. It is recommended that the patient determines failure of the intended step secondary to extremis or complaint, rather than the investigator. Supplemental oxygen is only acceptable in step 1 of the PDA. Patients on chronic supplemental oxygen therapy should be excluded. Further refinement of the scale might incorporate measures of oxygen saturation as an objective means to support a patient’s self-assessment, and perhaps respiration rate. Finally, the duration tolerated of each stage might also be factored in by subtracting a time penalty for patients unable to complete 3 min.

Using these raw data, treatment effects can be computed in a number of ways that make clinical sense for patients, nurses, and doctors. Patients want to get better quickly and then stay better. This can be computed as the rate of improvement over the first few hours or by the area-under-the-curve, as previously discussed. However, it should be noted that if the period of assessment is prolonged, even a striking initial benefit might be diluted if both groups of patients eventually get symptom relief. For this reason, as a general rule, dyspnoea assessment as an outcome for Stage A and B trials should be limited to the first 24 h after initial intervention.
In regards to patient related outcome instruments, it has been recommended that they be ‘based upon a clear conceptual framework, have evidence supporting content validity, and acceptable psychometric qualities.’ While the DSS is based upon a conceptual framework, the DSS is only a proposal and requires both methodological and statistical validation. This might be accomplished by including at least one previously utilized measure of dyspnoea, such as a VAS or Likert scale used in previous studies to provide comparative data. A ‘minimally important difference’ needs to be established to facilitate interpretation of results, which may vary depending on the setting. This difference would be refined over time as the scoring system evolves.

Proposal summary
First, for consideration as an outcome in an AHFS clinical trial, dyspnoea must be secondary to AHFS. An objective requirement, such as elevated BNP, is recommended to support this assumption. Other agreed upon clinical signs (rales, leg edema, JVD, S3) and/or radiographic evidence may also be important for appropriate patient selection.

Second, timing of dyspnoea measurement is critical. Measuring dyspnoea at its peak in order to measure the greatest differences from worst state would be ideal. Although most patients present with dyspnoea, it improves with standard therapy and thus delay in measurement may influence patient selection and limit the opportunity to demonstrate efficacy of novel therapies. Regardless, the timing of interventions and timing of measurements should be reported. This will depend upon the purpose and design of the clinical trial.

Finally and most importantly, the method of dyspnoea assessment should be standardized with future work directed towards determining the most valid assessment method, including validation of a dyspnoea measurement scale. We have proposed one method, based on the PDA, resulting in a Dyspnoea Severity Score (DSS) (Table 4). Comparison of this new measure to previously used assessments will also be important to evaluate its utility.

Conclusion
Dyspnoea is the most common presentation in AHFS, and is an important target for clinicians, investigators, and regulatory agencies. The absence of a valid tool with which to measure dyspnoea and the lack of standardization of patients’ assessment conditions and timing of measurement have hindered the progress. We propose a scoring system, the Dyspnoea Severity Score, which should provide a more detailed and consistent assessment of dyspnoea and is also better able to show changes over time and with treatment. This proposal requires validation. Presently, trials continue to target dyspnoea despite the absence of a validated standard and thus uniformity of measurement is needed.

Table 4 Dyspnoea severity score

<table>
<thead>
<tr>
<th>PDA (if successfully completed, continue to next stage)*</th>
<th>5-point Likert scale (Asked of patient at end of each successfully completed stage of the PDA)</th>
<th>DSS (Select number corresponding to Likert assessment on patient’s last completed stage of the PDA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sitting upright (&gt;60°) with supplemental oxygen (minimum 2L NC) (assessment after 3 min equilibration)</td>
<td>Worst possible shortness of breath</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Severely short of breath</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Moderately short of breath</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>Mildly short of breath</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Not at all short of breath</td>
<td>5</td>
</tr>
<tr>
<td>Sitting upright (&gt;60°), no oxygen (assessment after 3 min equilibration)</td>
<td>Worst possible shortness of breath</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Severely short of breath</td>
<td>7</td>
</tr>
<tr>
<td></td>
<td>Moderately short of breath</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Mildly short of breath</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Not at all short of breath</td>
<td>10</td>
</tr>
<tr>
<td>Supine (&lt;20° head elevation), no oxygen (assessment after 3 min equilibration)</td>
<td>Worst possible shortness of breath</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Severely short of breath</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Moderately short of breath</td>
<td>13</td>
</tr>
<tr>
<td></td>
<td>Mildly short of breath</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Not at all short of breath</td>
<td>15</td>
</tr>
<tr>
<td>Walking 50 m as fast as possible (post walk assessment)</td>
<td>Worst possible shortness of breath</td>
<td>16</td>
</tr>
<tr>
<td></td>
<td>Severely short of breath</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>Moderately short of breath</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Mildly short of breath</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Not at all short of breath</td>
<td>20</td>
</tr>
<tr>
<td>Six minute walk test (post-6 min walk assessment)</td>
<td>Worst possible shortness of breath</td>
<td>21</td>
</tr>
<tr>
<td></td>
<td>Severely short of breath</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>Moderately short of breath</td>
<td>23</td>
</tr>
<tr>
<td></td>
<td>Mildly short of breath</td>
<td>24</td>
</tr>
<tr>
<td></td>
<td>Not at all short of breath</td>
<td>25</td>
</tr>
</tbody>
</table>

* In general, patients who report moderate, mild or no breathlessness should proceed to the next stage.
Acknowledgements

The authors would like to thank Norman Stockbridge, MD for his comments and insight regarding this manuscript.

Conflict of interest: P.S.P. reported that he has received research support from PDL BioPharma and honoraria from Otsuka and Solvay Pharmaceuticals within the past 5 years. He is a consultant for PDL BioPharma.

J.G.F.C. reported research support from GlaxoSmithKline, Roche, AstraZeneca, Pfizer, Sanofi, Servier and the British Heart Foundation. He has also received speaker's honoraria and funding from Takeda.

J.R.T. reported that he has received research grants from Abbott Laboratories, Actelion Pharmaceuticals, Bristol-Myers Squibb, CardioDynamics, CHF Solutions, Cytokinetiks, GlaxoSmithKline, Merck, Myogen, NovaCardia, Sanofi-Aventis, Scios, Inc., the National Institutes of Health and the Department of Veterans Affairs. In the past 5 years, he has received honoraria from and/ or has been/is a consultant for Abbott Laboratories, Actelion Pharmaceuticals, Bristol-Myers Squibb, CardioDynamics, CHF Solutions, Cytokinetiks, Duke Clinical Research Institute, Medtronic, Merck, Momentum Research, Myogen, NovaCardia, Protein Design Labs, Scios, Inc., and Wyeth.

S.P.C. reported that he has received research grants/support from Biosignetics, Inovise Medical Inc., and the NIH/NHLBI. He has received honoraria from Otsuka Pharmaceuticals, Inovise Medical Inc., Scios, Abbott Point-of-Care, and PDL BioPharma.

G.S. reported that “This article was co-written by Dr. George Sopko in his private capacity. The views expressed in the article do not necessarily represent the views of the National Institutes of Health, the Department of Health and Human Services, nor the United States.

W.F.P. reported that he has received research grants/support from Abbott, Accumetrics, Biosite, CHF Solutions, Inovise, Inverness, Scios, and the Medicines Company. He has been on the speakers bureau and/or received honoraria from Abbott, Beckman-Coulter, Biosite, Inovise, Inverness, Ortho Clinical Diagnostics, PDL BioPharma, Vital Sensors, and Zeeland Pharma. He also reports a significant ownership interest in Vital Sensors.

G.C.F. reported that he has received research grants from Amgen, Biosite, Bristol-Myers Squibb, Boston Scientific, GlaxoSmithKline, Medtronic, Merck, Pfizer, Sanofi-Aventis, Scios, Inc., and the National Institutes of Health. He is/has been on the speaker’s bureau or has received honoraria in the past 5 years from Amgen, AstraZeneca, Biosite, Bristol-Myers Squibb, Boston Scientific, GlaxoSmithKline, King, Medtronic, Merck, NitroMed, Pfizer, Sanofi-Aventis, Schering-Plough, Scios, Inc., St. Jude Medical, and Wyeth. He is or has been a consultant for AstraZeneca, Amgen, Biosite, Bristol-Myers Squibb, Boston Scientific, CHF Solutions, GlaxoSmithKline, Medtronic, Merck, NitroMed, Pfizer, Sanofi-Aventis, Schering-Plough, Scios, Inc., and Wyeth.

A.Z.A. reported no conflicts of interest or disclosures.

J.D.K. reported research support and/or has served as a consultant and/or received honoraria from Schering Inc., Sanofi-Aventis, Accumetrics, Biosite Inc., Inovise Medical, PDL BioPharma, and Scios Inc.

A.B.S. reported that he has received research grants/support from Inverness Medical Solutions, Biosignetics, PDL BioPharma, Abbott Diagnostics, and the NIH/NHLBI (1RO1HL088459-01). He is or has been a consultant to Molecular Insight Pharmaceuticals, Abbott Diagnostics, and the Medicines Company.

M.T. reported no conflicts of interest or disclosures.

A.M. reported being a consultant for Abbott, Orion Pharma, Protein Design Biopharma, and SigmaTau. He has received honoraria from Abbott, Guidant, and Edwards Life Sciences.

E.R. reported no conflicts of interest or disclosures.

G.F. reported having received research support from PDL BioPharma, Medtronic, Otsuka, and Solvay Pharmaceuticals.

B.M. reported receiving consultation fees from Novacardia, Scios, PDL BioPharma, Cytokinetiks and research grants from Novacardia, Bristol Myers Squibb, and Sanofi-Aventis.

A.M. reported receiving research support from Biosite, Bayer, Abbott, and Roche and is a consultant for Biosite.

M.K. reported he serves on the executive committee for the ASCEND-HF trial sponsored by Johnson & Johnson.

M.G. reported receiving research grants from the NIH, Otsuka, SigmaTau, Merck, and Scios Inc.; he is or has been a consultant to Dobbio Pharm, ErreKappa Terapeutici, GlaxoSmithKline, Medtronic, Otsuka, Protein Design Lab, Scios Inc., and SigmaTau.

References


