In acute heart failure syndrome (AHFS), dyspnea is one of the most common but least understood presenting symptoms for hospitalization. For this reason, dyspnea relief is increasingly becoming a focus in the development of therapies for the treatment of AHFS, and currently stands as an acceptable primary end point for regulatory approval by governmental agencies. This raises the question of how best to measure such a subjective symptom. In this review, we will describe the basis for dyspnea, provide a detailed description of the strengths and weaknesses of the current best tools used to measure it, and describe future directions for future development of dyspnea measurement in AHFS. (Am Heart J 2010;160:209-14.)

Breathing discomfort, and its varying degrees of severity, is the most disturbing symptom patients with an acute heart failure syndrome (AHFS) can experience; and it often serves as the impetus to seek medical care. Acute heart failure syndrome is collectively defined as a gradual or rapid change in heart failure (HF) signs and symptoms resulting in a need for urgent therapy.1 This same sensation of breathlessness is what also drives patients with asthma and chronic obstructive pulmonary disease (COPD) to seek medical attention, and it would be helpful to describe the pathophysiology of dyspnea in AHFS. Dyspnea, as defined by the American Thoracic Society in their consensus on the mechanisms, diagnosis, and treatment of dyspnea, is “a term used to characterize 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 behavioural responses”.2 Although this is an agreed upon definition of the symptom, it is experienced differently by every patient and depending on the etiology. Patients with congestive HF will describe their dyspnea as “suffocating at rest” or “air hunger” or express the quality of rapid breathing rather than describe an increase in work of breathing that is commonly seen with pulmonary disease (ie, asthma, COPD).3,5

Pathophysiology

The pathophysiology is theorized to result from a patient’s perceived mismatch or dissociation between the efferent motor activity of the respiratory center in the brain and the incoming afferent signals from mechanical receptors in the airways, lungs, chest wall structures, and chemoreceptors in the blood (Table I).2,6 Thus, it has less to do with the status of intrinsic respiratory function and more to do with the unresolved and disjointed interpretation of information within the controls of the respiratory system.3 That is not to say that physiologic factors are spectators and not integral components. It has clearly been documented that the burdens of advanced age, malnutrition, anemia, and cardiopulmonary disease including congestive HF will initiate a cyclical and deleterious cascade of events that disrupts respiratory muscle function leading to a ventilatory challenge the system is unfit to meet, which further deteriorates respiratory function.4 If these factors could be modulated, then perhaps a better outcome could be achieved. Likewise, there is also an effort to identify the area of the cortex that processes information related to dyspnea with the goal of identifying a pathway that could be interrupted to prevent the uncomfortable sensation; however, it remains unidentified as evidenced by the lack of a cortical lesion that abolishes the sensation of dyspnea or a cortical area that causes it when stimulated.4

Evaluation of dyspnea

Because the understanding of the pathophysiology of dyspnea is limited as well as the technology to determine it, the best current measurements of dyspnea involve using quality of life measurements.2 These instruments can be divided into 3 categories based on how they assess dyspnea during activities of daily living, during exercise, and on the overall impact on health status (Table II).7,8

From the aDuke Clinical Research Institute, Duke University School of Medicine, Durham, NC, bCleveland Clinic, Cleveland, OH, and cDuke Translational Medicine Institute, Duke University School of Medicine, Durham, NC.
Submitted April 18, 2010; accepted May 15, 2010.
Reprint requests: Adrian F. Hernandez, MD, MHS, Duke Clinical Research Institute, PO Box 17969, Durham, NC 27715.
E-mail: adrian.hernandez@duke.edu
0002-8703/$ - see front matter
© 2010, Mosby, Inc. All rights reserved.
doi:10.1016/j.ahj.2010.05.020
Quality of life measurements have been used for years to measure qualities such as pain, anxiety, and stress that could not otherwise be directly quantified. These instruments have also been shown to be valid and reliable, meaning they have both the ability to measure a patient’s dyspnea and the quality of reproducible measurements. Currently, these are considered objective measurements of the subjective symptom of dyspnea; and because they come directly from the patient, they are clinically relevant to therapy management.

The first category of quality of life measurement involves using scales in the chronic setting with activities of daily living as the benchmark for degrees of dyspnea. These are the Medical Research Council (MRC) Dyspnoea Scale, the Oxygen Cost Diagram (OCD), the Baseline and Transition Dyspnea Indexes (BDI/TDI), and the University of California San Diego Shortness of Breath Questionnaire (UCSD SOBQ).

The MRC Dyspnoea Scale is the work of Sir Walter Morley Fletcher and the MRC in Wales in the 1940s. It was created in response to the problem of quantifying breathlessness in Welsh coal mine workers suffering from pneumoconiosis. Patients are asked to rate their degree of dyspnea on a scale from 1 to 5, with 1 being “not troubled by breathlessness except on strenuous exercise,” 2 being “short of breath when hurrying on the level or walking up a slight hill,” 3 being “walks slower than most people on the level, stops after a mile or so, or stops after 15 minutes walking at own pace,” 4 being “stops for breath after walking about 100 yards or after a few minutes on level ground,” and 5 being “too breathless to leave the house, or breathless when undressing.” This scale has been validated for use in COPD patients.

The OCD asks patients to rate their level of dyspnea corresponding to the oxygen requirements of 13 different activities ranked in ascending order from 0 to 100 according to the number of calories expended in performing these activities and represented as a value along a vertical 100-mm line. Sleeping, sitting, and standing as less calorie intense and therefore less oxygen demanding activities are ranked closer to 0, while walking, briskly or not, uphill is ranked as 100. Patients are asked to mark the point at which they believe they are when they are at their best. The score is tabulated as the distance from 0 in millimeters. A score of 100 noted no impairment at all. The main strength of this instrument is in its use as a description of a patient’s perceived exercise tolerance—it does not correlate well with objective changes in exercise tolerance. The overwhelming limitation of the OCD is that not all dyspneic patients can carry out the breadth of activities listed on the diagram. The frame of reference of the people incapable of performing all of the activities nullifies the widespread implementation of this particular instrument.

The BDI was developed to characterize the degree of activity that provokes dyspnea, the magnitude of effort necessary to carry out an activity, and the functional limitations in work and activities of daily living. The questionnaires were conceived for use in respiratory assessment; therefore, they are usually administered by health care providers familiar with history taking in respiratory disease. A cumulative grade is assigned to the patient’s baselines status and is based on the individual scores of the categories of functional impairment, magnitude of task, and magnitude of effort, which are assigned a grade from 0 to 4 (0 being significant impairment and 4 being no impairment). A cumulative grade closer to 0 corresponds to more severe impairment. The BDI is used in tandem with the TDI, which tracks changes from baseline. The same open-ended questionnaires were used with the same categories, but changes are logged on a scale from −3 (significant deterioration) to +3 (significant improvement). Overall, a cumulative grade from −9 to +9 is produced for changes from baseline, with a score closer to −9 marking a more significant deterioration. Although this

### Table I. Components of pathophysiology of dyspnea

<table>
<thead>
<tr>
<th>Components</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Afferent signals</td>
<td>Mechanical receptors in the airways, lungs, chest wall structures, and chemoreceptors in the blood</td>
</tr>
<tr>
<td>Efferent signals</td>
<td>Efferent motor activity of the respiratory center in the brain descending to the diaphragm, and accessory respiratory muscles</td>
</tr>
<tr>
<td>Central processing</td>
<td>Perceived mismatch or dissociation between afferent sensation and efferent motor</td>
</tr>
<tr>
<td>Physiologic factors</td>
<td>Intrinsic dysfunction of the respiratory system caused by the burden of cardiac, pulmonary, or cardiopulmonary disease</td>
</tr>
</tbody>
</table>

### Table II. Dyspnea measurement tools

<table>
<thead>
<tr>
<th>Category</th>
<th>Name of instrument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate dyspnea using scales in the chronic setting with ADLs as the benchmark for degrees of dyspnea</td>
<td>MRC Dyspnoea Scale</td>
<td>OCD, BDI, TDI, UCSD SOBQ</td>
</tr>
<tr>
<td>Rate level of dyspnea during cardiopulmonary exercise testing</td>
<td>Modified Borg scale</td>
<td>UCSD SOBQ</td>
</tr>
<tr>
<td>Rate the impact of dyspnea on the overall well-being of a patient</td>
<td>SGRQ, CRQ, CHFQ</td>
<td></td>
</tr>
</tbody>
</table>

ADLs, Activities of daily living.
instrument has been demonstrated to be valid and reliable and sensitive to changes in dyspnea levels in patients with respiratory disease, it has shortcomings in assessing HF patients. The major weaknesses with this instrument are that the questions asked by interviewers are not standardized and timely administration of the questionnaire requires some proficiency in its use. This instrument is very user dependent; therefore, significant interinterviewer variability can occur depending on the experience of the health care provider administering the questionnaire. To reduce variability when used in a clinical trial, the same interviewer would have to conduct every interview. In addition, the instrument has not been validated for use in assessment of dyspnea secondary to HF.

The UCSD SOBQ consists of 21 questions about the severity of dyspnea associated with activities of daily living and 3 questions about the extent of limitations in these activities caused by the dyspnea itself or the fear of dyspnea on an average day during the week leading up to answering these questions. Each question is rated from 0 (no breathlessness) to 5 (unable to complete a particular activity of daily living because of shortness of breath), producing an overall score from 0 to 120. The main weakness in using this instrument to measure dyspnea in HF patients is that patients are asked to rate their level of dyspnea with respect to certain activities they may no longer perform. In addition, it has not been proven to be sensitive enough to changes that take place in less than a week.

The second category involves the Borg scale, which gauges the level of absolute dyspnea by asking patients to rate their level of dyspnea during cardiopulmonary exercise testing. The original Borg scale was from 6 to 20, but the American College of Sports Medicine modified it to a scale from 0 to 10. A numerical score of 0 corresponds to a verbal qualifier of “no perceived dyspnea” after testing, whereas a score of 10 is considered “maximal” perceived dyspnea. In a study that measured expiratory flow and orthopnea in left ventricular HF, the Borg scale was shown to accurately measure dyspnea in both the seated and supine positions before and after treatment with vasodilators and diuretics until hospital discharge. The main weakness of the Borg scale in HF patients is that most of these patients would be incapable of performing the necessary cardiopulmonary testing when acutely hospitalized with AHFS.

The final category involves using question inventories that rate the impact of dyspnea on the overall well-being of a patient. These are the St. George’s Respiratory Questionnaire (SGRQ), the Chronic Respiratory Disease Questionnaire (CRQ), and the Chronic Heart Failure Questionnaire (CHFQ).

The SGRQ measures respiratory symptoms and activities that potentiate or are limited by dyspnea as well the overall impact of dyspnea on health status. The questionnaire asks 76 questions across the categories of symptoms (frequency and severity), activity (activities that cause or are limited by breathlessness), and impact (social functioning, psychological disturbances resulting from airways disease). The symptom category is rated on a 5-point Likert scale, and the activity and impact categories are yes/no responses. Each section is scored and weighted according to empirical data to produce a cumulative score from 0 to 100, with a higher score indicating worse health. The CRQ and CHFQ take a similar approach to assessing the impact of disease on quality of life. The CRQ and CHFQ questionnaires differ by only a single question and assess the disease limitations of activities of daily living using a 7-point scale (1 being extremely short of breath, 7 being not at all short of breath) to measure the domains of dyspnea, emotional function, mastery, and fatigue. The emotional function, mastery, and fatigue domains have standardized questions. The dyspnea domain is rated using the 7-point scale across 5 activities the patient has selected from memory or suggestion as being most important to their daily living. All of these questionnaires have been shown to accurately quantify the levels of dyspnea, with the scores on the questionnaires even demonstrating correlation to physiologic data yielded from pulmonary function tests. However, their main weakness, which prevents their use in clinical trials for HF patients, is that they are not sensitive enough to track the changes in dyspnea that can occur in a patient during their hospital stay.

**Dyspnea scales in AHFS clinical trials**

The most basic of requirements that new drugs must achieve to receive approval by the regulatory agencies is the demonstration that they improve either symptoms or clinical outcomes. The heterogeneity of symptoms, characteristics, and presentations of AHFS has limited the ability for creation of end points that satisfy the needs of the clinical community as well as regulatory agencies such as the Food and Drug Administration or European Medicines Agency. Because dyspnea is of the most common and disturbing experiences an AHFS patient encounters and because it satisfies one of the approved requirements, its relief has been targeted as a clinical end point. However, there is no current standardization to the measurement of dyspnea. And unfortunately, most of the aforementioned quality of life measurements have been validated for use in patients with chronic dyspnea secondary to pulmonary disease such as COPD or cystic fibrosis, or in lung patients undergoing pulmonary rehabilitation; thus, their application for use in the acute setting with AHFS patients not capable of exercise is limited at best.
although these instruments have been validated and proven reliable, they have not been demonstrated, for the most part, to be sensitive enough to track changes in dyspnea in HF patients over their average length of hospital stay. Furthermore, many of the questionnaires are time-consuming endeavors even for health care providers seasoned in their use.

All of these factor into the rationale of why the Visual Analog Scale (VAS) and Likert scales have been the most widely used and accepted measures of dyspnea in AHFS patients (Table III). Likert scales consist of 3-, 5-, or 7-point scales that ask patients to rate their level of improvement in response to therapy on a categorical spectrum ranging from markedly better to markedly worse or an appropriate variation. Moreover, the VAS asks patients to rate their level of breathing difficulty on a vertical numerical continuum with 0 at the bottom and 100 at the top, with 100 being the best imaginable ability to breathe and 0 being the worst conceivable dyspnea. The Likert scales and VAS have been established in multiple AHFS clinical trials such as VERITAS, RITZ-1/RITZ-2, VMAC, EVEREST, SURVIVE, and REVIVE-II as being valid and reliable instruments capable of discriminating the degree of a patient's dyspnea (Table III).

Minimal clinically important difference

Although the VAS and Likert scores have proven themselves to be the best tools among the quality of life measurements in measuring dyspnea in HF, they too suffer from shortcomings that can limit their use. For example, intersubject comparisons of VAS scores are hard to make because the maximum and minimum levels of breathlessness can be different for each individual—one person’s 50 is only another person’s 20. Although these scores will never be the same for everyone, the changes in perceived dyspnea scores before and after treatment in HF are most important and deserve more attention. The minimal clinically important difference (MCID) is “the smallest difference between scores in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient's management.”

The MCID in dyspnea scores in HF has not yet been fully explored.

The MCID for VAS has been investigated in 2 trials. In one prospective, observational study, the MCID in dyspnea was evaluated by assessing 156 patients before and after they received initial asthma therapy in an emergency department. During reassessment, subjects were asked to describe their asthma symptoms as “much better,” “a little better,” “no change,” “a little worse,” or “much worse.” The “mean VAS change among the ‘a little better’ subjects was 2.2 cm (95% CI 1.1, 3.4) which was significantly greater than the ‘−0.4 cm (95% CI −2.1, 1.4) change in the ‘unimproved’ subjects.”

Thus, a change of 2.2 cm or (22 mm) was found to be the minimal clinically significant improvement in VAS dyspnea scores. In another prospective observational study, 79 patients with diagnosed HF were asked to rate their level of dyspnea on a VAS before and after they received therapy. The study found that patients who had a higher recorded VAS score also had a significantly greater change in VAS. However, for all patients, the mean for a meaningful change in VAS was 21.1 mm (or 2.11 cm) (95% CI 12.3-29.9 mm). In essence, these studies, although well conducted, represent a first step in defining the MCID in dyspnea that HF patients experience as measured on VAS in response to therapeutic intervention. The consensus for the MCID appears to be between 21.1 and 22 mm. Moving forward, large randomized controlled trials are needed to form a more substantiated MCID.

The MCID has also been established in the CHFQ, TDI, and UCSD SOBQ. The CHFQ was established in a retrospective study that compared the results of 3 previous studies that used the CHFQ. Patients were asked in the CHFQ to rate their shortness of breath during day-to-day activities, their level of fatigue, and how they were feeling emotionally and then compared that with how they had improved overall on a 15-point global rating scale that ranged from −7 (a great deal worse),
through 0 (no change), to +7 (a great deal better). A global rating scale change from either −3 to −1 or +1 to +3 corresponded to a significant decrease in dyspnea. This corresponded to a mean change in 3 points per question in the dyspnea domain of the CHFQ, which averaged out to be 0.5 point per question within each domain. The MCID for the TDI was established with a multinational clinical trial of 997 patients with COPD. In this study, they found that a mean change of 1 unit in the TDI focal score corresponded to a clinically significant decrease in dyspnea. The MCID for the UCSD SOBQ was established in a study of 164 chronic lung disease patients before and after pulmonary rehabilitation. A mean change of 5 units corresponded to a clinically significant decrease in dyspnea.

Future methods for dyspnea evaluation

For any drug to receive US Food and Drug Administration approval, it has to demonstrate in a clinical trial either a decrease in mortality or relief of symptoms associated with a particular disease or condition. For this reason, the outcome end points for the development of therapies for AHFS have included decreased mortality, decreased hospital stay and rehospitalization, decreased use of special interventions, and relief of symptoms, as well as combinations of some or all of these. Among the relief of symptoms, dyspnea is the most important to patients. The improvement a patient reports in the assessment of his or her own dyspnea is one of the most important standards by which efficacy of therapy in AHFS is ascertained. Moreover, dyspnea relief serves as a viable reflection for physicians of patient improvement in the short term. Immediate relief of dyspnea mere hours after presentation can lead to more rapid stabilization of patients that can theoretically be discharged with a reduced length of stay. Consequently, this affects data collection for other clinical end points, underscoring the importance of furthering the evolution of dyspnea relief as a clinical end point in the treatment of AHFS. Despite the crux of improvement in a patient’s clinical course being placed on dyspnea relief and the development of new AHFS therapies being based on it also, the measurement of dyspnea itself has not been well defined in the clinical setting and remains the next priority in AHFS end points. A concerted effort is needed to accurately gauge the dyspnea continuum and its MCID to allow for accurate tracking of quantifiable changes in response to therapy that will help guide the development of new therapies for this burdensome disease.

One potential direction is development of a combination of methods that relate to dyspnea. These new approaches must for instance take into account the potential effect of a patient’s position on dyspnea assessment. This issue was explored in the recently published URGENT-dyspnoea study. In this trial, HF patients were initiated on the standard of care in European medical centers (ie, intravenous diuretics) and underwent dyspnea assessment 6 hours after initiating therapy. The majority of patients were evaluated in the sitting position, but those with less severe dyspnea in this position were graduated to dyspnea assessment in the supine position. It was found that orthopnea may be refractory to treatment in the acute setting, as patients evaluated in the supine position reported less improvement in their dyspnea than their upright counterparts. With these and multiple other considerations in mind, the Dyspnea Severity Score (DSS) has been developed as a way to standardize dyspnea measurements. The DSS consists of asking patients to rate their level of dyspnea on a 5-point Likert scale in each category of the Provocative Dyspnea Assessment, which has patients sitting upright with oxygen, sitting upright without oxygen, lying supine without oxygen, walking 50 m as fast as possible, and a post–6-minute walk test. The DSS ranges from 1 to 25 and essentially measures when patients can no longer progress in performance. Although the DSS does well to incorporate objective measures, its overall scoring is still entirely reliant on patient reporting. It incorporates no concrete objective data and, as a result, is subject to the variation that is inherent in most other dyspnea assessment tools. The DSS, although quantifiable, still lacks a tangible MCID and is tedious and hence challenging to ascertain in large clinical trials. A mega clinical trial (ASCEND-HF) assessing dyspnea relief in AHFS is under way and may advance our understanding of pathophysiologic correlates of dyspnea relief. Dyspnea will be measured using the 7-point Likert scale in all patients at 6 and 24 hours after initiation of therapy. Change in weight, urine volume, biomarkers including natriuretic peptides (in a subset), and a respiratory substudy measuring peak expiratory flow rate will provide additional data to help ascertain an MCID in these patients. As promising as the DSS is, it has not been validated for use in any clinical trials to date.

Conclusions

Dyspnea is a complex pathophysiologic state that is not well understood and is deeply disturbing to patients who suffer from it. The best efforts to measure dyspnea to provide a basis upon which clinical trials for the development of new therapies for AHFS can be conducted or patients’ improvement can be clinically judged are aimed at using quality of life measurements. Among these quality of life measurements, the Likert scale and VAS have been established as being the best combination of valid, reliable, and easy to use instruments for measuring dyspnea in the clinical setting; with the DSS being the first promising, yet untested, step in
standardizing dyspnea assessment. Quantifying the exact significance in degree of change in dyspnea with these quality of life measurements needs future attention; however, initial steps have been taken by exploring the MCID for the VAS, in particular. In improving the evaluation of dyspnea relief as a benchmark for AHFS intervention efficacy, additional steps may be needed. To date, as dictated by the understanding of the pathophysiology, dyspnea measurements have relied almost entirely on subjective data from either the patient or the health care provider. Future studies may consider incorporating objective data in addition to subjective measures, although symptom relief is at the heart of the problem from a patient’s perspective.

References
2248-58.