Development and Validation of a Patient Questionnaire to Determine New York Heart Association Classification

SPENCER H. KUBO, MD,[1] STEVEN SCHULMAN, MD,[2] RANDALL C. STARLING, MD, MPH,[3] MARIELL JESSUP, MD,[4] DEBORAH WENTWORTH, MPH,[1] AND DANIEL BURKHOFF, MD, PhD[5]

St. Paul, Minnesota; Baltimore, Maryland; Cleveland, Ohio; Philadelphia, Pennsylvania; New York, New York

ABSTRACT

Background: New York Heart Association (NYHA) classification correlates with quality of life and is useful in tracking changes in status in clinical trials. We explored methods to determine NYHA class in multicenter trials where double-blind conditions could not be maintained.

Methods and results: A questionnaire was developed containing 7 major questions based on the standard definitions of NYHA classes. The questionnaire was administered to 116 patients with varying degrees of heart failure. When comparing NYHA determined by physicians at the site to NYHA assigned by 3 independent physician graders, there was an approximate 60% concordance. Concordance between independent reviewers was approximately 75%. Results of repeat grading of 30 randomly selected questionnaires indicated that graders provided the same score 90% of the time. Thus, although there were some differences from the site determination of NYHA class, the questionnaire had good inter- and intragrader reproducibility. In a second group of 103 patients enrolled in an ongoing device-intervention trial, we demonstrated that it is feasible to employ the questionnaire in a multicenter trial. Finally, NYHA class was correlated with quality of life and peak exercise oxygen consumption.

Conclusions: A standardized questionnaire provided an approximate 60% concordance in assigning NYHA classification compared to the site assessment with approximately 90% reproducibility. This approach may be useful to determine NYHA classification within the context of clinical trials where blinded conditions are not possible.

Key Words: New York Heart Association, heart failure, functional class, questionnaires.

The New York Heart Association (NYHA) classification[1] is a 4-point semiquantitative index of functional status of patients with heart failure. NYHA class is widely accepted and useful clinically because it correlates with quality of life[2–4] and survival.[5] When measured serially over time, it provides a means of tracking disease progression and response to therapeutic interventions. Although it is "subjective," NYHA class has also been used in many clinical trials as a pivotal demonstration of efficacy for both pharmacologic and device interventions.

In clinical trials, however, it is generally accepted that double-blind conditions must be employed for NYHA determinations to be a valid assessment, in order to avoid bias. However, double-blind conditions are not possible in many device and surgical intervention trials because the implanter or surgeon cannot be blinded and it is not ethically justifiable to perform sham surgery in control patients. Some device trials have used additional blinded investigators at the site to make clinical assessments. However, concerns have been raised about inadvertent unblinding resulting from casual conversations or exposure to certain laboratory data.

Furthermore, although most physicians are experienced in assigning a NYHA class, the method of assignment is not standardized. To our knowledge, reproducibility and consistency of determining NYHA class have never been established. Goldman et al.[6] developed a specific activity scale in which patient functional class was based on the estimated

From the [1]Acorn Cardiovascular Inc., St. Paul, Minnesota; [2]Department of Medicine, Johns Hopkins University, Baltimore, Maryland; [3]Department of Cardiovascular Medicine, Kaufman Center for Heart Failure, Cleveland University Foundation, Cleveland, Ohio; [4]Department of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania; [5]Department of Medicine, Columbia University, New York, New York.

Manuscript received June 30, 2003; revised manuscript received September 22, 2003; revised manuscript accepted October 16, 2003.

Reprint requests: Spencer H. Kubo, MD, Acorn Cardiovascular, Inc., 601 Campus Drive, St. Paul, MN 55112.

This study was supported in part by Acorn Cardiovascular, Inc.
1071-9164/$ - see front matter

doi:10.1016/j.cardfail.2003.10.005

metabolic cost of different activities. However, the output of the specific activity scale was not exactly analogous to NYHA and many of the queries did not appear to be relevant to a contemporary population.

We therefore set out to develop standardized methods of NYHA classification that could be suitable for use in multicenter trials, particularly where double-blind conditions could not be maintained. We hypothesized that a centralized "core lab" could effectively isolate the assessor from patient contact and inadvertent unblinding at the site. However, a centralized core lab requires a standardized set of patient responses. Two of us (SS and DB) had prior experience with a core lab and a standardized assessment of Canadian Cardiovascular Society Angina Score, which was required within the context of an unblinded study of a device treatment for angina.[7]

Accordingly, we created a questionnaire with the intent that it would mimic the types of questions that a physician might ask during a patient interview to determine NYHA class. The questionnaire had to be comprehensive, because a centralized core lab would not have the opportunity to examine the patient, ask follow-up questions, or look at test results. The questionnaire was then tested and validated in a group of 116 patients. The questionnaires were scored by 3 experienced clinicians who served as independent blinded core lab graders. Statistical methods were used to test intra- and intergrader reproducibility and correlations between NYHA scores assigned by the site physicians and those provided by the core lab. Second, correlations between core lab assigned NYHA class and several measures of functional capacity (quality of life, maximal exercise capacity, and submaximal exercise performance) were completed to assess construct validity. Finally, the questionnaire was applied to a multicenter trial to assess the feasibility of implementation across many centers.

The primary objective of this study was to develop a survey that could determine NYHA class by a third-party core lab in a manner that could be utilized in a nonblinded multicenter trial. It was not meant to replace or improve the traditional method by which clinicians assess NYHA in everyday clinical encounters.

Methods

Questionnaire Development

The standard definitions of the 4 NYHA classes are summarized in Table 1. The activities and levels of exertion indicated in these definitions were translated into a set of questions with possible patient responses. These questions were designed to mimic a conversation between a physician and patient using the following principles:

1. The questionnaire was designed to be completed by a person with a minimal amount of training (eg, a person without a medical background) who reads a script and elicits standardized responses from a patient without coaching or biasing responses. Accordingly, the questionnaire should be simple and self-explanatory.
2. Allowable patient responses to an inquiry about whether he or she participates in a specific activity covered a range of frequencies (never, rarely, sometimes, or frequently) rather than simple yes/no answers.
3. Questions were designed to detect patient avoidance of activities that produced symptoms. For example, a patient may indicate that he or she never has symptoms with a specific activity merely because the patient avoided performing the activity because of the anticipation of symptoms.
4. Open-ended questions, although useful in a clinical setting, were considered impractical for a survey.
5. A global health status question (2) using a numeric scale ranging from normal health (10) to severe debilitation (1) was considered useful as a means of providing internal validation of overall functional status obtained from the standardized set of questions.

Based on the these considerations, a questionnaire with 7 major questions was developed (Fig. 1). The questionnaire was reviewed with a small group of heart failure physician specialists to refine wording and eliminate ambiguous phrases prior to use with patients.

The concept of using an automated algorithm for assigning NYHA from patient responses was not considered in the present study because of the inability of simple algorithms to reconcile inconsistent patient responses. Instead, a separate scoring tool was created that would group the frequency of activities, the frequency of symptoms associated with these activities, and the presence or absence of rest symptoms. This worksheet provided the framework to interpret symptoms and derive NYHA classification.

Validation

The questionnaire was applied to 116 heart failure patients at 4 outpatient heart failure clinics; these patients are designated as the core lab validation patients. The questionnaire was generally administered by a nurse or physician assistant by reading the scripted instructions. The responses were recorded on the questionnaire. The primary heart failure physician at the heart failure clinic independently provided a NYHA classification based on his or her normal interaction with the patient. Most patients were seen in the course of routine clinical follow-up and were not participating in any clinical study, which could have biased the physicians' assessment.

Table 1. Definition of the New York Heart Association classifications

Class I. Patients with cardiac disease but without resulting limitation of physical activity. Ordinary physical activity does not cause undue fatigue, palpitation, dyspnea, or anginal pain.
Class II. Patients with cardiac disease resulting in slight limitation of physical activity. They are comfortable at rest. Ordinary physical activity results in fatigue, palpitation, dyspnea, or anginal pain.
Class III. Patients with cardiac disease resulting in marked limitation of physical activity. They are comfortable at rest. Less than ordinary activity causes fatigue, palpitation, dyspnea, or anginal pain.
Class IV. Patients with cardiac disease resulting in inability to carry on any physical activity without discomfort. Symptoms of heart failure or the anginal syndrome may be present even at rest. If any physical activity is undertaken, discomfort increases.

AHA Science Advisory Assessment of functional capacity in clinical and research applications, #71-0190. Circulation 2000;102:1591-7.

Completed questionnaires were then sent to three physicians who each served as an independent, blinded "core lab graders." These 3 graders read each questionnaire and used the scoring tools as a worksheet to derive a NYHA classification.

Correlations Between Core Lab NYHA and Measures of Functional Capacity

To determine the feasibility of the questionnaire in a multicenter trial, and to investigate the relationship of core lab assigned NYHA class to other measures of quality of life and functional capacity, the questionnaire was administered to patients enrolled in an ongoing multicenter trial. The baseline data from this second group (103 patients) are designated as the functional capacity correlation patients. The Acorn Pivotal Clinical Trial is a multicenter (29 site) randomized controlled assessment of the safety and efficacy of the Acorn CorCap™ Cardiac Support Device. The device is a novel textile mesh jacket that is implanted around the heart and designed to reduce diastolic wall stress and promote reverse remodeling of the heart. The study enrolled 300 patients with heart failure stratified into 1 of 2 arms according to the presence or absence of a clinical indication for mitral valve repair or replacement (MVR). The MVR stratum included patients who had a clinical indication for mitral valve or replacement because of >3 + mitral regurgitation. The Cardiac Support Device–only stratum included patients with heart failure who had <2 + mitral regurgitation and did not have an indication for mitral valve surgery. Patients had to be NYHA class III/IV (as determined by the site investigator) to enter into both strata; patients with NYHA class II symptoms could additionally be entered into the MVR stratum. Baseline evaluations also included completion of a Minnesota Living with Heart Failure Questionnaire (MLHFQ), a 6-minute hall walk test, and a maximal exercise test using metabolic gas analysis for the measurement of peak exercise oxygen consumption (PVO_2), all using standardized protocols.[2,8,9]

Each of 29 sites identified a clinician (physician or nurse) not directly involved in the clinical trial and who was blinded to treatment assignment, to administer the questionnaire. The clinician received training instructions on how to interview the patient and how to record the answers on the form. Surveys were coded to mask patient identification and were sent to 1 core lab reviewer for NYHA assignment; these were then compared with the baseline assessments of MLHFQ, 6MWT, and PVO_2. Because the trial was ongoing when the questionnaire was developed, baseline data were available only for 103 patients of the 300 that were enrolled in the trial.

Statistical Analysis

Contingency tables were produced from data obtained from core lab validation patients comparing the results of the site assigned NYHA with those assigned by each of the 3 reviewers. The reviewer's scores were also compared with each other. The agreement (percent of time the assessed NYHA exactly matched) and 95% confidence intervals were computed for each of these comparisons.

To adjust for the degree of agreement that might be expected by chance alone, a weighted kappa statistic was computed.[10] Weighted kappa adjusts for chance agreement and ranges from −1 meaning no agreement to +1 indicating perfect agreement. A weighted kappa of 0 indicates that the degree of agreement is equivalent to the agreement expected by chance alone.

Data obtained from the functional capacity correlation patients including MLHFQ, PVO_2, and 6MWT were compared with the core lab determination of NYHA using Spearman rank correlation coefficients. $P < .05$ was considered statistically significant.

Results

Core Lab Validation

The validation study included a total of 116 patients. Based on the site physician's assessments, there were 23 NYHA class I patients, 43 NYHA class II patients, 39 NYHA class III patients, and 11 NYHA class IV patients. Table 2 shows the agreement between NYHA class assigned by the site physician and each of the 3 independent NYHA core lab graders. Numbers on the diagonal of each matrix indicate the number of concordant grades between the site physicians and the blinded reviewers.

For reviewer #1 (Table 2a) there was concordance with the site in 65% of the cases (75/116 cases). There were 29 overestimates (sum of cases above the diagonal line) and 12 underestimates (sum of cases below the diagonal line). Thus, on average, this grader had an approximate 15% [= (29−12)/116)*100] bias to assign a higher NYHA class than the site physicians. The weighted kappa for this comparison was 0.63, indicating a moderately high concordance between the site and Reviewer #1 even after correcting for random chance agreement. Reviewer #2 showed a slightly lower overall concordance with site assessments (Table 2b, 57% concordance, weighted kappa 0.55). There were roughly equal number of entries above and below the diagonal (26 and 24, respectively), so there was only a 2% [=(26−24)/116*100] overall overgrading bias compared with the site. Reviewer #3 (Table 2c) was similar to #2 in having 59% concordance and 0.56 weighted kappa. However, this reviewer had a 24% overgrading bias [=(38−10)/116] with regard to site assessments.

As summarized in Table 3, there was less intergrader variability than grader-site variability. Reviewers 1 and 2 had a 72% concordance (weighted kappa 0.74), with a 14% overgrading bias by Reviewer 1. Reviewers 2 and 3 also had a 72% concordance (weighted kappa 0.73), with a 24% overgrading bias by Reviewer 3. Finally, Reviewers 1 and 3 had a 78% grading concordance (weighted kappa 0.78), with a 12% overgrading bias by Reviewer 3.

To investigate the reproducibility of the scoring system, 30 randomly selected questionnaires were reviewed a second time by 2 of the reviewers. This second review occurred 3 months after the original review. Patient identities were removed. For both reviewers, grades were identical in 27 of the 30 cases (90% agreement, weighted kappa 0.88). The deviant scores differed from the original scores by only 1 class. For each reviewer, 2 of the scores were higher than the original score and 1 was lower than the original score.

Finally, detailed analysis showed that these results did not differ significantly when data from each site were analyzed individually.

Functional Correlations

Table 4 contains the mean values of functional test results obtained from 103 patients included in the functional

CENTRAL ASSESSMENT OF NYHA FUNCTIONAL CLASS

Note: All Subscript & Shaded Areas Are for Data Entry Purposes Only.

PATIENT INITIALS: (F)_____(M)_____(L)_____

Note: This interview must be conducted by a clinician who is blinded to the treatment assignment.

Please Read to Patient:

"I will be asking you a few questions about your health status. Please respond to the best of your ability with the answer that you feel best describes your health. Do not tell me whether or not you have had mitral valve surgery or whether or not you have received the Acorn Cardiac Support Device. All answers need to be yours and based on your own experience." (*Please ask clinical staff, relatives and other visitors to leave the room if they are with you).*

Questions can be answered as Yes/No or with one of 4 responses:

Never or none of the time

Rarely, a little of the time, or on occasion

Some of the time, a moderate amount of the time

Frequently, most of the time

Answer all applicable questions.

1. a. How often do you walk up and down stairs? (8-12 steps)

$_1$☐ Never → b. Do you avoid stairs because it makes you tired or short of breath?
$_0$☐ No $_1$☐ Yes **Go to #2.**

$_2$☐ Rarely
$_3$☐ Sometimes
$_4$☐ Frequently

c. How often do you get short of breath or tired when you walk up or down a flight of stairs at a normal pace under normal conditions?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently

d. How often do you get short of breath or tired when you walk up or down a flight of stairs quickly?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #2.**

2. a. How often do you engage in strenuous work or prolonged exertion at work or play?

$_1$☐ Never → b. Do you think you would get short of breath or tired if you engaged in these activities?
$_0$☐ No ☐ Yes **Go to #3.**

$_2$☐ Rarely
$_3$☐ Sometimes
$_4$☐ Frequently

c. How often do you get short of breath or tired doing strenuous work or prolonged exertion?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #3.**

Fig. 1. NYHA questionnaire.

3. a. How often do you go for walks, either outside or inside, on level ground at a normal pace under normal conditions?

$_1$☐ Never → b. Do you avoid walks because it will make you short of breath or tired?
$_0$☐ No $_1$☐ Yes **Go to #4.**

$_2$☐ Rarely
$_3$☐ Sometimes → c. How often would you get short of breath or tired if you walked less than 1 block?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently
$_4$☐ Frequently

d. How often would you get short of breath or tired if you walked more than 2 blocks?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently

e. How often do you find yourself walking more slowly than usual?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #4.**

4. a. How often do you get short of breath or tired when you are sitting doing nothing or when you are sleeping?

$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #5.**

5. a. How often do you walk up hills?

$_1$☐ Never → b. Do you avoid walking up hills because it makes you short of breath or tired?
$_0$☐ No $_1$☐ Yes **Go to #6.**

$_2$☐ Rarely
$_3$☐ Sometimes → c. How often do you get short of breath or tired walking up hills in normal weather?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #6.**
$_4$☐ Frequently

6. a. How often do you go out in the cold/windy or hot/humid weather?

$_1$☐ Never → b. Do you avoid cold/windy or hot/humid weather because it makes you short of breath or tired?
$_0$☐ No $_1$☐ Yes **Go to #7.**

$_2$☐ Rarely
$_3$☐ Sometimes → c. How often do you get short of breath or tired when you go outside in the cold, windy or hot, humid weather?
$_1$☐ Never $_2$☐ Rarely $_3$☐ Sometimes $_4$☐ Frequently **Go to #7.**
$_4$☐ Frequently

7. a. On a scale of 1-10, with 10 being perfectly normal and 1 being near death, how would you rate yourself?

1	2	3	4	5	6	7	8	9	10

Fig. 1. (Continued)

capacity correlation patients to test the construct validity of the questionnaire. Patients are grouped according to the core lab–assigned NYHA class. Mean values for MLHFQ, PVO_2 consumption and 6MWT results varied incrementally as NYHA class increased. NYHA was positively correlated with MLWHF ($r = .47$; $P < .0001$) and negatively correlated with 6MWT ($r = -.32$, $P = .001$). The correlation of NYHA and PVO_2, however, was not significant ($r = -.15$, $P = .20$). These findings are consistent with other studies that show that a gradient from NYHA I to IV in terms of quality of life, maximal exercise capacity, and submaximal exercise performance.[2,4,8,11]

Discussion

The main objective of the present study was to develop, validate, and assess the reproducibility of a questionnaire that could be used by a core laboratory to assign NYHA classification within the context of multicenter clinical trials. The questionnaire was simple and easy to complete. The information was sufficient for a core lab grader (an experienced clinician) to make a NYHA determination that was reasonably close to the NYHA determination made by the clinician at the site, with approximately 60% concordance. More importantly, there was even greater concordance (approximately 75%) among 3 independent core lab graders, and a high degree of reproducibility (approximately 90%), when the same reviewer evaluated surveys a second time. Core lab NYHA determinations were correlated with other measures of functional capacity, including quality of life and 6MWT. Finally, the feasibility of using the questionnaire and core lab within the context of a multicenter trial has been demonstrated.

It is important to note that the development of this questionnaire was not intended to replace or even validate the traditional method by which NYHA class is determined by clinicians. NYHA class is the gold standard and has been used successfully in many double-blind clinical trials as an efficacy end point. This survey was developed to fill a void in clinical trials in which one needs to assess NYHA but under conditions where blinding is not possible.

That the core lab graders did not agree perfectly with the site determination of NYHA was anticipated. Site clinicians have many pieces of information available to them that can be used in determining NYHA class, such as PVO_2 consumption, ejection fraction, and hemodynamics. It can

Table 2. Comparison of site assessment to reviewers, all patients reviewed (n = 116)*

Site × Reviewer 1
Reviewer 1

Site	1	2	3	4
1	15	7	1	0
2	3	27	12	1
3	0	5	26	8
4	0	0	4	7

n = 116
Agreement = 75; 65% (95% CI = 56%, 73%)
Weighted kappa = .63

Site × Reviewer 2
Reviewer 2

Site	1	2	3	4
1	18	5	0	0
2	11	19	10	3
3	2	6	23	8
4	0	1	4	6

n = 116
Agreement = 66; 57% (95% CI = 48%, 66%)
Weighted kappa = .55

Site × Reviewer 3
Reviewer 3

Site	1	2	3	4
1	13	8	2	0
2	2	27	11	3
3	0	3	22	14
4	0	0	5	6

n = 116
Agreement = 68; 59% (95% CI = 50%, 68%)
Weighted kappa = .56

*Excludes patients identified as having incomplete data.

Table 3. Comparison of reviewers to each other, all patients reviewed (n = 116)*

Reviewer 1 × Reviewer 2
Reviewer 2

Reviewer 1	1	2	3	4
1	17	1	0	0
2	13	23	3	0
3	1	7	31	4
4	0	0	3	13

n = 116
Agreement = 84; 72% (95% CI = 64%, 81%)
Weighted kappa = .74

Reviewer 2 × Reviewer 3
Reviewer 3

Reviewer 2	1	2	3	4
1	15	15	1	0
2	0	22	8	1
3	0	1	30	6
4	0	0	1	16

n = 116
Agreement = 83; 72% (95% CI = 63%, 80%)
Weighted kappa = .73

Reviewer 1 × Reviewer 3
Reviewer 3

Reviewer 1	1	2	3	4
1	14	4	0	0
2	1	31	7	0
3	0	3	31	9
4	0	0	2	14

n = 116
Agreement = 90; 78% (95% CI = 70%, 85%)
Weighted kappa = .78

*Excludes patients identified as having incomplete data.

Table 4. Construct validity*

Core laboratory NYHA assessment	MLHFQ	PVO_2	6MWT
I	14 (n = 1)	18.3 (n = 1)	392 (n = 1)
II	34 (n = 7)	16.2 (n = 4)	361 (n = 7)
III	57 (n = 42)	14.6 (n = 23)	371 (n = 41)
IV	69 (n = 53)	14.0 (n = 33)	324 (n = 53)

6MWT, 6-minute hall walk test (meters); MLHFQ, Minnesota Living with Heart Failure Questionnaire (units); NYHA, New York Heart Association; PVO_2, peak exercise oxygen consumption (mL/kg/min).

*Number of patients in each cell indicated in parentheses.

be argued that this additional information should not be used in determining functional class, but, in reality, it is synthesized into the NYHA determination all the time. Moreover, to our knowledge, NYHA class assignment as performed in clinical practice has never been standardized, formally validated, nor tested for reproducibility. Therefore, it is not possible to determine if the site or the core lab determination was more "accurate" or a "better" description of NYHA class. Both are "accurate" insofar that they reflect the information available at that time.

The key difference between the traditional method and the method described here is that the questionnaire-based method uses a standard set of questions. We are not aware of prior studies and did not attempt to measure the consistency and reproducibility of NYHA determinations performed in the traditional manner but suspect that there could be great variability given the lack of standardization of NYHA assessment in routine clinical practice. The questionnaire-based method showed good internal consistency and reproducibility among the different reviewers. It has been suggested that the consistency and reproducibility of the measure might be an especially important feature when using NYHA to detect treatment differences in a clinical trial.

Although simple and reproducible, there are a number of important limitations of the current questionnaire. It does not specifically ask the patient for any value judgments related to the impact of their disease on their quality of life. Further, it is not a comprehensive measurement of health status or functional capacity, so that it would not replace the need for other patient reported instruments, such as the MLHFQ, the Kansas City Cardiomyopathy Questionnaire, or the SF-36 (short form-36). It is simply an alternate but more systematic, reliable, and reproducible means of determining NYHA class that can be used in multicenter trials.

Clinicians may fundamentally question the need for such a standard questionnaire for NYHA determination when they can easily determine NYHA in any given patient. This new questionnaire is not intended to replace the traditional method of determining NYHA in routine clinical practice. Despite the universal awareness and acceptance of the traditional means of assessing NYHA, there are no standardized methods for performing or even training clinicians. If different clinicians ask different questions or use different data sources (eg, exercise testing) in making such assessments, there are potential hazards when NYHA determinations are compared across clinics. Therefore, even though the questionnaire-based method requires the completion of a short form, the standardization and rigor is expected to enhance the consistency and reproducibility of NYHA determinations.

The current questionnaire requires a grading by an experienced clinician to determine NYHA. It may be possible, however, to employ neutral networks or an artificial intelligence algorithm on the expanded data set to develop computer generated "decision rules." Therefore, future versions of the questionnaire might eliminate the experienced clinician grader and instead use an automated NYHA determination after the questions are answered.

Symptoms of angina and heart failure frequently coexist in patients with heart failure and are difficult to reliably distinguish. Further, myocardial ischemia could cause both symptoms of angina and heart failure. The current questionnaire cannot distinguish angina and heart failure and even experienced clinicians may have difficulties in this task. However, both processes could be expected to produce symptomatic limitations that could be measured in a global NYHA assessment.

It must be emphasized, that within the context of an unblinded study, the approach described herein does not eliminate potential bias on the part of the patient, because knowledge of treatment group can still influence his or her responses to questions related to well being (ie, the placebo effect).

The questionnaire developed in the present study proved to be simple to administer with essentially no training and relatively easy to score by experienced reviewers. The results obtained from a single grader correlated well with several well-established parameters of functional capacity. Based upon the present analysis, this questionnaire and core lab based strategy are currently being used in a clinical trial for a cardiac support device (Acorn Cardiovascular, Inc.). Results from that 300-patient study will provide additional data on the degree of agreement between the clinician assessment and a standardized method of NYHA determination and will help further define the utility of this approach.

References

1. AHA Science Advisory. Assessment of functional capacity in clinical and research applications. Circulation 2000;102:1591–7.
2. Rector TS, Kubo SH, Cohn JN. Patient's self assessment of their congestive heart failure: Part 2: content, reliability and validity of a new measure, The Minnesota Living with Heart Failure Questionnaire. Heart Fail 1987;3:198–209.
3. Al-Kaade S, Hauptman PJ. Health-related quality of life measurement in heart failure: challenges for the new millenium. J Card Failure 2001;7:194–201.
4. Green CP, Porter CB, Bresnahan DR, Spertus JA. Development and evaluation of the Kansas City Cardiomyopathy Questionnaire: a new health status measure for heart failure. J Am Coll Cardiol 2000;35: 1245–55.
5. Francis GS, Kubo SH. Prognostic factors affecting diagnosis and treatment of congestive heart failure. Curr Prob Cardio1 1989;14:631–71.

6. Goldman L, Hashimoto B, Look EF, Loscalzo A. Comparative reproducibility and validity of systems for assessing cardiovascular functional class: advantages of a new specific activity scale. Circulation 1981;69:1227–34.
7. Oesterle SN, Sanborn TA, Ali N, Resar J, Ramee SR, Heuser R, et al. Percutaneous transmyocardial laser revascularization for severe angina: the PACIFIC Randomized Trial. Lancet 2000;356: 1705–10.
8. Guyatt GH, Sullivan MJ, Thompson PJ, Fallen EL, Pugsley SO, Taylor DW, et al. The 6-minute walk: a new measure of exercise in patients with chronic heart failure. Can Med Assoc J 1985;132: 919–23.
9. Wasserman K. Principles of exercise testing and interpretation. 3rd ed. Baltimore: Lippincott Williams & Wilkins, 1999.
10. Fleiss JL, Cohn J. The equivalence of weighted kappa and the intra class correlation coefficient as measures of reliability. Educ Psych Measure 1973;33:613–9.
11. Szlachcic J, Massie BM, Kramer BL, Topic N, Tubau J. Correlates and prognostic implications of exercise capacity in chronic congestive heart failure. Am J Cardiol 1985;5:1037–42.