Original research

Revisiting heart failure assessment based on objective measures in NYHA functional classes I and II

Mariana Blacher,[1,2] André Zimerman,[1,3] Pedro H B Engster,[3] Eduardo Grespan,[3] Carisi A Polanczyk,[1,2,3] Marciane M Rover,[4] José A de Figueiredo Neto,[5] Luiz C Danzmann,[6] Eduardo G Bertoldi,[7] Marcus Vinicius Simões,[8] Luis Beck-da-Silva,[1,2,3] Andréia Biolo,[1,2,3] Luis E Rohde [1,2,3]

► Additional material is published online only. To view, please visit the journal online (http://dx.doi.org/10.1136/heartjnl-2020-317984).

For numbered affiliations see end of article.

Correspondence to
Dr Luis E Rohde, Hospital de Clinicas de Porto Alegre, Porto Alegre, Brazil;
rohde.le@gmail.com

MB and AZ contributed equally.

Received 7 August 2020
Revised 11 November 2020
Accepted 16 November 2020

To cite: Blacher M, Zimerman A, Engster PHB, *et al*. *Heart* Epub ahead of print: [*please include* Day Month Year]. doi:10.1136/heartjnl-2020-317984

ABSTRACT

Objective New York Heart Association (NYHA) functional class plays a central role in heart failure (HF) assessment but might be unreliable in mild presentations. We compared objective measures of HF functional evaluation between patients classified as NYHA I and II in the *Rede Brasileira de Estudos em Insuficiência Cardíaca* (ReBIC)-1 Trial.

Methods The ReBIC-1 Trial included outpatients with stable HF with reduced ejection fraction. All patients had simultaneous protocol-defined assessment of NYHA class, 6 min walk test (6MWT), N-terminal pro-brain natriuretic peptide (NT-proBNP) levels and patient's self-perception of dyspnoea using a Visual Analogue Scale (VAS, range 0–100).

Results Of 188 included patients with HF, 122 (65%) were classified as NYHA I and 66 (35%) as NYHA II at baseline. Although NYHA class I patients had lower dyspnoea VAS Scores (median 16 (IQR, 4–30) for class I vs 27.5 (11–49) for class II, p=0.001), overlap between classes was substantial (density overlap=60%). A similar profile was observed for NT-proBNP levels (620 pg/mL (248–1333) vs 778 (421–1737), p=0.015; overlap=78%) and for 6MWT distance (400 m (330–466) vs 351 m (286–408), p=0.028; overlap=64%). Among NYHA class I patients, 19%–34% had one marker of HF severity (VAS Score >30 points, 6MWT <300 m or NT-proBNP levels >1000 pg/mL) and 6%–10% had two of them. Temporal change in functional class was not accompanied by variation on dyspnoea VAS (p=0.14).

Conclusions Most patients classified as NYHA classes I and II had similar self-perception of their limitation, objective physical capabilities and levels of natriuretic peptides. These results suggest the NYHA classification poorly discriminates patients with mild HF.

INTRODUCTION

Heart failure (HF) is a prevalent and morbid syndrome with a heterogeneous clinical presentation.[1 2] The comprehensive evaluation of patients with HF is challenging and involves assessing both heart structure and functional status.[3–5] For patients with HF with reduced ejection fraction, the foremost discrimination relies on the New York Heart Association (NYHA) classification, that sorts patients with HF into four classes according to their aptitude to perform physical activities.[6] Originally described in 1928, this classification has become a key inclusion criterion in HF trials[7–9] and a central feature of international clinical guidelines.

An important but rather neglected aspect of NYHA functional class assessment is its subjectivity. In particular, the distinction between asymptomatic (NYHA class I) and minimally symptomatic (NYHA class II) patients can be difficult as functional status can fluctuate in short intervals, depends on how the patient perceives his symptoms, on how much he is willing to expose and on each physicians' individual perception. It is not surprising that interobserver and intraobserver reproducibility in NYHA evaluation are suboptimal.[10–13] Despite limitations of this subjective assessment, a myriad of treatments—including dapagliflozin, sacubitril-valsartan and implantable cardioverter-defibrillators—is offered to patients classified as NYHA II, but not to those considered NYHA I.

Such important therapeutic implications operate under the assumption that NYHA classification adequately identifies patients who can benefit from each offered treatment. In this scenario, there are scarce contemporary data objectively evaluating the capacity of NYHA assessment in discriminating between patients with HF. In particular, few reports have considered the patients' own perception of his physical limitations.[14] In this report, we use data from the '*Rede Brasileira de Estudos em Insuficiência Cardíaca*' (ReBIC)-1 Trial[15] to describe the association between NYHA class and three markers of disease severity: the 6 min walk test (6MWT), N-terminal pro-brain natriuretic peptide (NT-ProBNP) levels and self-assessed dyspnoea using a Visual Analogue Scale (VAS).

METHODS

Patient population

We analysed patients with HF enrolled in the ReBIC-1 Trial, a randomised, double blind, placebo-controlled trial that evaluated the safety and tolerability of discontinuing furosemide in outpatients with chronic stable HF and no evidence of congestion. The detailed protocol and final results have been published.[15 16] In brief, 188 outpatients with

HF were enrolled at 11 Brazilian clinical sites and randomised to furosemide withdrawal (n=95) or maintenance (n=93). Adult patients with HF were enrolled in the original study if they fulfilled the following criteria: NYHA functional class I or II; left ventricular ejection fraction (LVEF) ≤45%; no HF-related hospitalisations or visits to the emergency room within 6 months before the screening visit; treatment with a stable dose of furosemide (40 mg or 80 mg per day) for at least 6 months before the screening visit; serum potassium <5 mmol/L; optimal HF treatment with ACE inhibitors (ACEi) or angiotensin receptor blockers (ARBs) and β-blockers, unless contraindicated or not tolerated. The study concluded that withdrawing furosemide did not lead to either increased self-perception of dyspnoea or increased need of furosemide reuse. All patients provided written informed consent before enrolment and randomisation.

Study logistics

Patients were evaluated at baseline (V1) and during follow-up visits at 15 (V2), 45 (V3) and 90 (V4) days after randomisation. At baseline and at the final visit all patients underwent assessment of (1) NYHA functional class, (2) A Clinical Congestion Score,[17 18] (3) Dyspnoea perception, (4) Standard 6MWT, and (5) NT-proBNP levels using total heparinised venous blood and a point-of-care equipment (*COBAS* h 232, measuring range 60–9000 pg/mL; F. Hoffmann-La Roche, Basel, Swiss). NYHA class and congestion were defined by the attending physician (a senior cardiologist), who was unaware of additional test scores. Dyspnoea was self-assessed using a VAS: patients were asked to mark their level of dyspnoea on a horizontal line based on their perception of shortness of breath during the prior week. The VAS ranged from 0 to 100, and independently and blindly reassessed by the coordination centre for all enrolled patients at all time points.

Statistical analysis

Data are expressed as mean (±SD), median (25th–75th centiles) or absolute number (percentage). NT-proBNP levels were normalised with logarithmic transformation. Dyspnoea VAS did not depict a normal distribution after logarithmic transformation and was analysed using non-parametrical statistical tests. For comparisons of NYHA functional classes with other measures (table 1), the unpaired t-test was used for normally distributed variables (6MWT and a logarithmic NT-proBNP) and the Wilcoxon rank-sum test for non-normally distributed variables (VAS). Kernel density estimations, with Silverman's rule-of-thumb bandwidth selection, were used to calculate the overlapping area for different measures between patients in NYHA classes I and II.[19] Comparisons between NYHA classes were displayed as violin plots. For temporal variations in NYHA class and dyspnoea VAS, NT-proBNP and 6MWT, we compared baseline versus final visits for all patients. Data were presented as scatter plots with a least squares regression line. Variation of dyspnoea VAS had a normal distribution and was analysed using analysis of variance. Cut-offs of HF severity depicted in table 2

Table 1 Clinical characteristics of the population

Characteristics‡	All	NYHA class I	NYHA class II
Patients, n (%)	188 (100)	122 (65)	66 (35)
Age, years	59.2±12.1	58.5±12.1	60.4±12.1
Female sex, n (%)	48 (25.5)	30 (24.6)	18 (27.3)
Hypertension, n (%)	117 (62.2)	73 (59.8)	44 (66.7)
Diabetes, n (%)	52 (27.7)	34 (27.9)	18 (27.3)
Atrial fibrillation, n (%)	24 (12.8)	18 (14.8)	6 (9.1)
LVEF, %	32.1±7.8	32.6±7.9	31.1±7.6
Ischaemic aetiology, n (%)	63 (33.5)	40 (32.8)	23 (34.8)
Hypertensive aetiology, n (%)	38 (20.2)	25 (20.5)	13 (19.7)
Systolic blood pressure, mm Hg	122±21.3	122±20.8	123±22.3
Weight, kg	76.9 (67.1–86.6)	76.0 (67.2–86.7)	77.8 (67.3–85.9)
Creatinine, mg/dL	1.1±0.4	1.1±0.3	1.2±0.5
Potassium, mEq/L	4.6±0.5	4.6±0.5	4.6±0.4
β-blocker, n (%)	187 (99.5)	121 (99.2)	66 (100)
ACEi or ARB, n (%)	171 (91.0)	114 (93.4)	57 (86.4)
Spironolactone, n (%)	134 (71.3)	88 (72.1)	46 (69.7)
Furosemide, n (%)			
40 mg	154 (81.9)	101 (82.8)	53 (80.3)
80 mg	34 (18.1)	21 (17.2)	13 (19.7)
Hydrochlorothiazide, n (%)	13 (6.9)	9 (7.4)	4 (6.1)
MAGGIC Risk Score*†	16 (11–19)	15 (10–18)	17 (14–21)
Dyspnoea VAS Score, mm†	20 (5–38)	16 (4–30)	27.5 (11–49)
NT-proBNP, pg/ml†	646 (284–1697)	620 (248–1333)	778 (421–1737)
NT-proBNP, pg/ml, $\log_{10}$†	2.82±0.53	2.75±0.56	2.94±0.46
6MWT, m†	377±110	390±112	352±104
CCS, points	2 (2–3)	2 (1–2)	3 (2–4)

*Missing values (1.1%) for the MAGGIC Risk Score (31) were computed with multiple imputation.
†p<0.05.
‡Continuous data are displayed as mean±SD or median (IQR).
ACEi, ACE inhibitor; ARB, angiotensin receptor blocker; CCS, Clinical Congestion Score; LVEF, left ventricular ejection fraction; MAGGIC, Meta-Analysis Global Group in Chronic Heart Failure; 6MWT, 6 min walk test; NT-proBNP, N-terminal pro-brain natriuretic peptide; NYHA, New York Heart Association; VAS, Visual Analogue Scale.

Table 2 NYHA functional class and cut-offs of dyspnoea self-assessment, exercise capacity and congestion

Parameter	All n=188	NYHA class I n=122	NYHA class II n=66	P value*
A. Dyspnoea VAS >30 points	60 (32)	30 (25)	29 (44)	0.01
B. 6MWT distance <300 m	41 (23)	22 (19)	19 (30)	0.11
C. NT-proBNP >1000 pg/mL	68 (36)	42 (34)	26 (39)	0.60
A and B	16 (9)	7 (6)	9 (14)	0.12
A and C	23 (12)	12 (10)	10 (15)	0.38
B and C	17 (9)	9 (7)	7 (11)	0.60
A and B and C	5 (3)	2 (2)	3 (5)	0.47
Neither	62 (33)	49 (40)	13 (20)	0.006

*χ^2 test between groups.
6MWT, 6 min walk test; NT-proBNP, N-terminal pro-brain natriuretic peptide; NYHA, New York Heart Association; VAS, Visual Analogue Scale.

were defined according to previously validated reports.[20–22] An a priori sample size calculation was not performed, because sample size was originally determined from the ReBIC-1 Trial. To assess predictors of NYHA classification and of self-assessed VAS Scores, we estimated a multivariable logistic regression and a linear regression, respectively, with centre-level random effects. Values of $p<0.05$ were considered statistically significant. All analyses were performed using R (V.3.6.1 or higher) and STATA software (V.16, StataCorp, College Station, Texas, USA).

Patient and public involvement

No patients were involved in developing the research question, study design, choice of outcome measures or recruitment in the primary trial.

RESULTS

Patient characteristics

Detailed clinical characteristics of the ReBIC-1 Trial patient population have been previously published.[16] Table 1 describes main clinical characteristics of the 188 patients with HF who were enrolled, stratified according to NYHA class at baseline. Mean age was 59±12 years, mean LVEF was 32%±8% (IQR 26%–39%), most patients were men and 65% were classified as NYHA class I. Baseline drug therapy was stable and optimised for most patients. Target doses were attained in 70% of users of ACEi/ARB, 63% of ß-blockers and 94% of mineralocorticoid antagonists. Mean and median baseline self-assessed dyspnoea VAS were 22.8 (±19) and 20 (5–38), respectively, ranging from 0 to 73. Mean and median baseline NT-proBNP levels were 1300 (±1660) pg/mL and 646 (284–1697) pg/mL, ranging from 25 pg/mL to 9000 pg/mL. Mean and median baseline 6MWT distance were 377 (±110) m and 380 (308-450) m, ranging from 48 m to 710 m.

NYHA class versus self-assessed dyspnoea, NT-proBNP levels and 6MWT

Figure 1 displays the three measures of HF status, stratified by baseline NYHA functional class. In a population level, the mean ranks (VAS) or mean values ($\log_{10}$ NT-proBNP and 6MWT) differed statistically between NYHA classes I and II. However, in an individual level, the overlap was substantial. For the self-assessed dyspnoea VAS Scale, overlap between patients from the two classes was 60%; for $\log_{10}$NT-proBNP values, 78%; and for the 6MWT distance, 64%.

NYHA class and severity cut-offs

Table 2 describes NYHA functional class according to cut-offs that indicate more severe HF in each of the following domains:

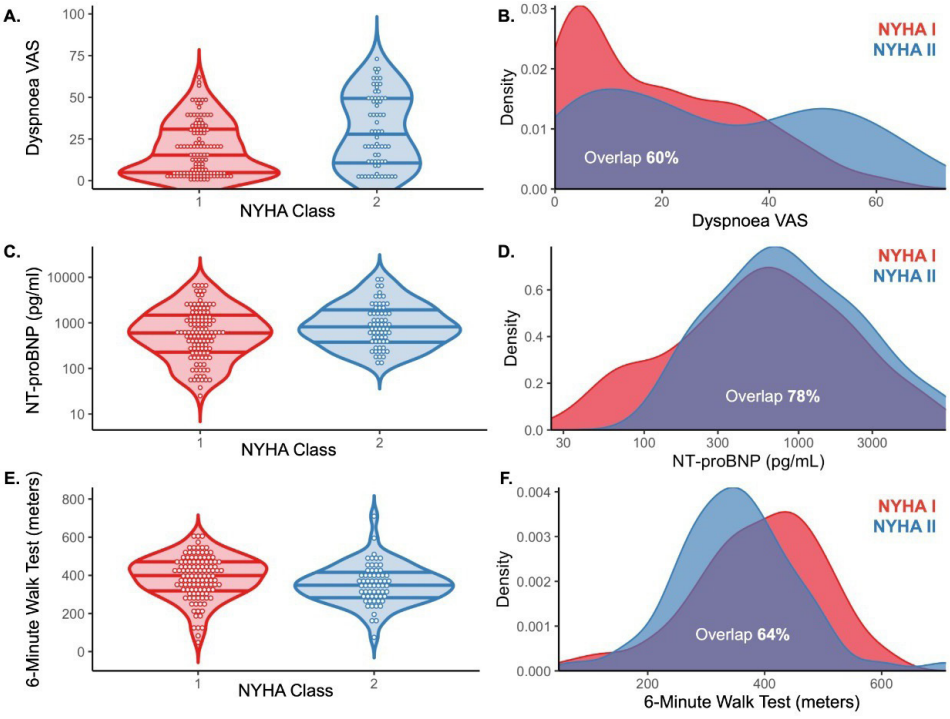

Figure 1 Violin plots with individual data on Visual Analogue Scales (VAS) of dyspnoea (A), N-terminal pro-brain natriuretic peptide (NT-proBNP) levels (C) and 6 min walk test (6MWT) distance (E) in New York Heart Association (NYHA) functional classes I and II. Middle horizontal lines represent median values and upper and lower lines represent the 75th and 25th centiles, respectively. Panels B, D and F illustrate density histograms for the same parameters.

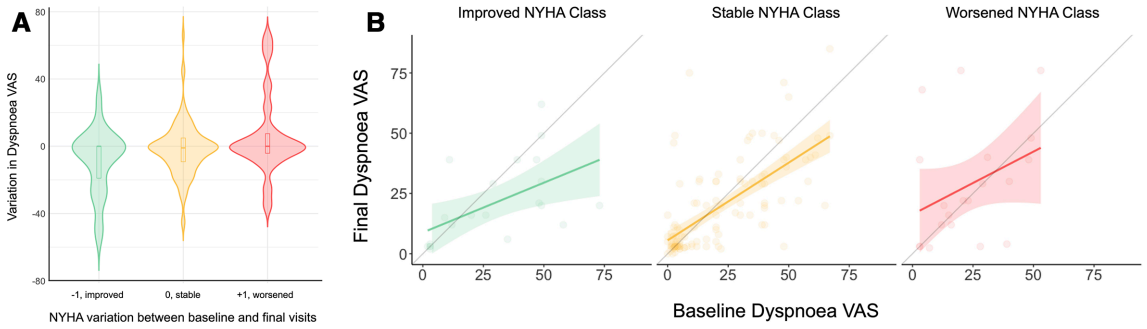

Figure 2 Violin plots (A) of the variation of Visual Analogue Scales (VAS) of dyspnoea according to improvement (−1), stability (0) or worsening (+1) in New York Heart Association (NYHA) functional class. Middle horizontal lines represent median values and upper and lower lines represent the 75th and 25th centiles, respectively. Panel B plots the first and the second VAS Scores according to improvement (−1), stability (0) or worsening (+1) in NYHA functional class. Regression lines and 95% CIs are also depicted.

dyspnoea (VAS Score >30 points), exercise capacity (6MWT distance <300 m) and ventricular distension (NT-proBNP >1000 pg/mL). Among patients in NYHA class I, 19%–34% had one of these markers, and 6%–10% had two of them simultaneously. Based on the assumption that a patient in NYHA class I would not have an important self-perception of dyspnoea (VAS Score >30 points), 30 (25%) patients of the current NYHA I sample would be reclassified as symptomatic HF. Moreover, 15 of 66 (23%) patients classified as NYHA II were virtually asymptomatic (VAS Score <10 points).

Variation of NYHA class

To better address the reliability of NYHA class assessment, we analysed whether temporal improvement, stability or worsening of functional class were associated with changes in self-assessed dyspnoea (figure 2A). For variations in NYHA category between baseline and final assessments, median variation in dyspnoea VAS was 0 (IQR for decreased class, −19 to 0; for increased class, −3 to 16), while median VAS change for stable NYHA class was −1 (IQR −9 to 5). We were unable to detect any statistical relationship between changes in NYHA Score and perception of dyspnoea (p value from Kruskal-Wallis rank-sum test=0.14). Figure 2B displays correlation lines between first and second dyspnoea VAS assessments, separated by change in functional class status. Neither improvement (left panel) nor worsening (right panel) of NYHA classes depicted the expected changes in dyspnoea VAS. Online supplemental figures 1 and 2 illustrate temporal variation of functional class related to 6MWT and NT-proBNP levels.

Variation on dyspnoea VAS

In an analysis stratified by the temporal variation of dyspnoea VAS (< −10 points, between −9 and 9, and >10 points; online supplemental figure 3), we also did not observe major reciprocal changes in NYHA functional class.

Predictors of NYHA class and VAS Scores

In an exploratory multivariable regression model to predict NYHA class, the only variable significantly associated with NYHA II (vs NYHA I) was the VAS Score (OR 1.43 for every 10-point increase). The discrimination accuracy of the model, which also included demographics, NT-proBNP and 6MWT, was limited (area under the curve of 0.72). Online supplemental figure 4 illustrates an exploratory multivariable linear regression model to assess predictors of VAS Scores.

DISCUSSION

NYHA functional class, a subjective definition, plays a central role in the clinical, therapeutic and prognostic assessment of HF.[7–9] We evaluated data from a contemporary clinical trial[15 16] to objectively characterise how patients classified as NYHA classes I and II differed in several clinical measures, evaluated both cross-sectionally and longitudinally. Although we observed significant differences among several markers of physical capacity according to NYHA class, the overall overlap of data was substantial, indicating that group-level differences cannot be readily translated to individual patients. Indeed, most patients classified by their physicians as classes I and II have similar perceptions of their own limitation, objective physical capabilities and levels of natriuretic peptides. Also, whenever changes in NYHA functional class occurred over time, we did not observe reciprocal changes in the perception of dyspnoea. These observations challenge the reliability of the NYHA classification for individual patients with mild HF, and consequently question each therapeutic indication whose benefit relies on the dichotomisation between NYHA classes I and II.

Previous studies have addressed the reliability of NYHA classification. In 1981, Goldman *et al* found that a second physician's NYHA assessment would reproduce the first classification in only 56%, demonstrating the subjectivity of NYHA classification and the need for other tools to assess HF functional status.[12] Caraballo *et al* objectively compared NYHA II and III participants of the HF-ACTION (Efficacy and Safety of Exercise Training in Patients With Chronic Heart Failure) and GUIDE-IT (Guiding Evidence-Based Therapy Using Biomarker Intensified Treatment in Heart Failure) Trials. Overlap between the two classes was substantial for every evaluated measure, including NT-proBNP (79% and 69%), 6MWT (63% and 54%) and LVEF (88% and 83%).[23] Notably, Yap *et al* found that the distance in the 6MWT discriminated changes in NYHA functional class in patients with moderate to severe HF (between classes II to IV), but not in mild forms of the syndrome (between classes I and II).[24] In this scenario, the 6MWT provides objective, reliable and valid data, and is a strong and independent predictor of long-term mortality in patients with HF.[25]

The relationship between functional class and natriuretic peptide levels has also been addressed previously. A retrospective analysis of the 'Diuretic Optimisation Strategy Evaluation in Acute Heart Failure (DOSE-AHF)' Trial assessed the relationship between markers of decongestion (weight loss, net fluid loss and per cent reduction in serum NT-proBNP level) and relief of symptoms as defined by the dyspnoea VAS 72 hours

after randomisation to different furosemide strategies. Interestingly, all markers were poorly correlated with dyspnoea relief, suggesting that the instruments available to measure dyspnoea level still have low sensitivity and poor validity, and demonstrating the gap in the clinical assessment of dyspnoea and congestion.[26] Taken together, these reports suggest that while individual markers are poorly correlated with HF severity, using a combination of instruments can improve discrimination. VAS Scores emerge as a suitable option to incorporate patient-centred measures for the assessment of dyspnoea. Although VAS Scores have been used mostly in acute decompensated HF as an indicator of dyspnoea relief,[27] they have also been validated in the outpatient setting to predict future cardiac events.[28] Interestingly, variations in dyspnoea VAS Scores were associated with changes in impedance cardiography parameters in patients with chronic HF, suggesting that it provides objective data reflecting changes in disease status.[29]

It is unquestionable that NYHA classification is one of many powerful prognostic predictors in HF, easily obtained from clinical history, capable of differentiating extremes of functional capacity.[12] Our results, however, reinforce the concept that it might be limited to discriminate more subtle changes or variations in functional status. We have demonstrated group-level statistically significant differences of small magnitude in the self-assessed dyspnoea VAS, 6MWT distance and levels of NT-proBNP between NYHA class I and class II. Interpretation of these results deserve careful consideration, as most individual patients with HF had substantial overlap in essential features related to their overall clinical status. The huge overlap between these important measures of functionality mitigate the clinical relevance of these statistical group differences. Even though the NYHA I group has a lower mean VAS Score, the individual patient with a lower than average VAS Score cannot be automatically classified as NYHA I. In addition, improvement or deterioration in NYHA functional class did not capture reliable reciprocal changes in the perception of dyspnoea. Explanations for these findings might be intrinsically related to both patients' and physicians' perceptions of HF symptoms, and to their willingness to inform and understand. All these facets might also be considerably influenced by cognitive, social, cultural, emotional and environmental aspects of the patient-physician relationship. While it is arguable that similarities between NYHA classes I and II may reflect poor patient communication or faulty medical clinical judgement, this further corroborates the limitations of the subjective NYHA functional class to reliably discriminate patients in different risk strata.

In daily practice, we face the clinical dilemma of determining whether a patient with HF is in NYHA functional class I. The clinical implications of such definitions should not be minimised. In contrast to other prognostic markers, the definition of NYHA class has itself major therapeutic implications. Patients with HF in NYHA functional class I might not be eligible for several disease-modifying drugs and device therapies based on current recommendations from most international HF guidelines.[3–5] As such, it is conceivable that a substantial subgroup of patients labelled as NYHA I might be currently denied important therapies that could be beneficial. In precision medicine, decision-making relies on individualised assessments, which require both objective and subjective measures of disease.

Limitations

Some aspects of our study design deserve consideration. Although the clinical definition of NYHA functional class has never been formally validated in different languages, it uses simple and standardised criteria that are routinely applied by doctors worldwide, including the trial's physicians. Our findings are compatible with previous studies, in different languages and cultures, that also demonstrated the inconsistency of this classification.[10 12] Another potential concern is the small sample size of patients with HF that were evaluated and the relatively short-term follow-up (90 days). Accordingly, larger studies with longer follow-up and in distinct scenarios are needed to corroborate and to validate the present results. Third, the study population enrolled in the ReBIC-1 Trial involved stable outpatients with mild forms of HF (only patients with NYHA classes I and II). Therefore, findings from the current analysis should not be extrapolated to other populations with more severe HF symptoms. Fourth, serial cardiopulmonary exercise testing is the gold standard to measure functional status in patients with HF, but this was unavailable as in most clinical and research settings.

In conclusion, our data suggest that the NYHA classification has limited validity for patients with mild HF, and many times will not correspond to the expected findings based on patient-reported dyspnoea or objective metrics of functionality and ventricular distension. One should be cautious when facing patients with HF labelled as NYHA classes I and II, and establishing clinical decisions solely based on this simple but limited assessment, as it might have important therapeutic implications.

Key messages

What is already known on this subject?

- New York Heart Association (NYHA) functional class plays a central role in heart failure (HF) assessment and represents a central feature to decide whether a patient is eligible for life-saving medications or devices.

What might this study add?

- We demonstrated that most patients classified by their physicians as NYHA classes I and II have similar perceptions of their own limitations, objective physical capabilities and levels of natriuretic peptides. In addition, whenever changes in NYHA functional class occurred over time, we did not observe reciprocal changes in objective parameters of functional status.

How might this impact on clinical practice?

- Our observations challenge the reliability of the NYHA classification for individual patients with mild HF, and consequently question clinical decisions that rely solely on the dichotomisation between NYHA classes I and II.

Author affiliations
[1]Post-Graduate Program in Cardiology and Cardiovascular Sciences, Faculdade de Medicina da Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil
[2]Cardiovascular Division, Hospital Moinhos de Vento, Porto Alegre, Brazil
[3]Cardiovascular Division, Hospital de Clínicas de Porto Alegre, Porto Alegre, Brazil
[4]Instituto de Cardiologia, Porto Alegre, Brazil
[5]Universidade Federal do Maranhão, Sao Luis, Maranhão, Brazil
[6]Universidade Luterana do Brasil, Canoas, Rio Grande do Sul, Brazil
[7]Universidade Federal de Pelotas, Pelotas, Brazil
[8]Faculdade de Medicina de Ribeirão Preto, Universidade de Sao Paulo, Sao Paulo, Brazil

Twitter André Zimerman @AndreZimerman

Contributors All authors contributed to the design and conduct of the study, data collection, analysis and preparation of the manuscript. LER supervised the study and is the guarantor.

Funding This study was supported in part by the 'Instituto de Avaliação em Tecnologia em Saúde (IATS)' and funded by the National Council for Scientific and Technological Development (CNPq, Brazil) with grants for the establishment of the Brazilian Research Network in Heart Failure (# 467110/2014-0). LER received a research fellowship from CNPq (# 30833/2017-1).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study was approved by the institutional review board at each site and was registered at ClinicalTrials.gov (NCT02689180).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement De-identified participant data are available upon reasonable request.

ORCID iD

Luis E Rohde http://orcid.org/0000-0002-6557-393X

REFERENCES

1 Ponikowski P, Anker SD, AlHabib KF, *et al*. Heart failure: preventing disease and death worldwide. *ESC Heart Fail* 2014;1:4–25.
2 AS G, Mozaffarian D, Roger VL, *et al*. Heart disease and stroke statistics — 2013 update a report from the American heart association. *Circulation* 2013;127:6–245.
3 Rohde LE, Montera MW, Bocchi EA, *et al*. Diretriz Brasileira de Insuficiência Cardíaca Crônica E Aguda. *Arq Bras Cardiol* 2018;111:436–539.
4 Ponikowski P, Voors AA, Anker SD, *et al*. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC)Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J* 2016;37:2129–200.
5 Yancy CW, Jessup M, Bozkurt B, *et al*. 2017 ACC/AHA/HFSA focused update of the 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology/American heart association Task force on clinical practice guidelines and the heart failure Society of America. *Circulation* 2017;136:e137–61.
6 The Criteria Committee of the New York Heart Association. *Nomenclature and Criteria for Diagnosis of Diseases of the Heart and Great Vessels*. 9th edn. Boston, Mass: Little, Brown & Co, 1994: 253–6.
7 McMurray JJV, Packer M, Desai AS, *et al*. Angiotensin-neprilysin inhibition versus enalapril in heart failure. *N Engl J Med* 2014;371:993–1004.
8 Stone GW, Lindenfeld J, Abraham WT, *et al*. Transcatheter mitral-valve repair in patients with heart failure. *N Engl J Med* 2018;379:2307–18.
9 Marrouche NF, Brachmann J, Andresen D, *et al*. Catheter ablation for atrial fibrillation with heart failure. *N Engl J Med* 2018;378:417–27.
10 Bennett JA, Riegel B, Bittner V, *et al*. Validity and reliability of the NYHA classes for measuring research outcomes in patients with cardiac disease. *Heart Lung* 2002;31:262–70.
11 Forgiarini LA, Rubleski A, Douglas G, LA FJ, Garcia D, *et al*. Evaluation of respiratory muscle strength and pulmonary function in heart failure patients. *Arq Bras Cardiol* 2007;89:36–41.
12 Goldman L, Hashimoto B, Cook EF, *et al*. Comparative reproducibility and validity of systems for assessing cardiovascular functional class: advantages of a new specific activity scale. *Circulation* 1981;64:1227–34.
13 Madsen BK, Hansen JF, Stokholm KH, *et al*. Chronic congestive heart failure. description and survival of 190 consecutive patients with a diagnosis of chronic congestive heart failure based on clinical signs and symptoms. *Eur Heart J* 1994;15:303–10.
14 Felker GM, Lee KL, Bull DA, *et al*. Diuretic strategies in patients with acute decompensated heart failure. *N Engl J Med* 2011;364:797–805.
15 da Rosa PR, Rohde LE, Doebber M, *et al*. Rational and design of a randomized, double-blind, multicenter trial to evaluate the safety and tolerability of furosemide withdrawal in stable chronic outpatients with heart failure: the ReBIC-1 trial. *Am Heart J* 2017;194:125–31.
16 Rohde LE, Rover MM, Figueiredo Neto JA, *et al*. Short-Term diuretic withdrawal in stable outpatients with mild heart failure and NO fluid retention receiving optimal therapy: a double-blind, multicentre, randomized trial. *Eur Heart J* 2019;40:3605–12.
17 Rohde LE, Beck-da-Silva L, Goldraich L, *et al*. Reliability and prognostic value of traditional signs and symptoms in outpatients with congestive heart failure. *Can J Cardiol* 2004;20:697–702.
18 Sauer J, Rabelo ER, Castro RA, *et al*. Nurses' performance in classifying heart failure patients based on physical exam: comparison with cardiologist's physical exam and levels of N-terminal pro-B-type natriuretic peptide. *J Clin Nurs* 2010;19:3381–9.
19 Silverman BW. *Density estimation for statistics and data analysis*. London: Chapman & Hall, 1986.
20 Mueller C, McDonald K, de Boer RA, *et al*. Heart failure association of the European Society of cardiology practical guidance on the use of natriuretic peptide concentrations. *Eur J Heart Fail* 2019;21:715–31.
21 Giannitsi S, Bougiakli M, Bechlioulis A, *et al*. 6-Minute walking test: a useful tool in the management of heart failure patients. *Ther Adv Cardiovasc Dis* 2019;13:1753944719870084.
22 Flynn KE, Lin L, Ellis SJ, *et al*. Outcomes, health policy, and managed care: relationships between patient-reported outcome measures and clinical measures in outpatients with heart failure. *Am Heart J* 2009;158:S64–71.
23 Caraballo C, Desai NR, Mulder H, *et al*. Clinical implications of the new York heart association classification. *J Am Heart Assoc* 2019;8:e014240.
24 Yap J, Lim FY, Gao F, *et al*. Correlation of the new York heart association classification and the 6-minute walk distance: a systematic review. *Clin Cardiol* 2015;38:621–8.
25 Rostagno C, Olivo G, Comeglio M, *et al*. Prognostic value of 6-minute walk corridor test in patients with mild to moderate heart failure: comparison with other methods of functional evaluation. *Eur J Heart Fail* 2003;5:247–52.
26 Kociol RD, McNulty SE, Hernandez AF, *et al*. Markers of decongestion, dyspnea relief, and clinical outcomes among patients hospitalized with acute heart failure. *Circ Heart Fail* 2013;6:240–5.
27 Mebazaa A, Pang PS, Tavares M, *et al*. The impact of early standard therapy on dyspnoea in patients with acute heart failure: the URGENT-dyspnoea study. *Eur Heart J* 2010;31:832–41.
28 Havranek EP, Lapuerta P, Simon TA, *et al*. A health perception score predicts cardiac events in patients with heart failure: results from the IMPRESS trial. *J Card Fail* 2001;7:153–7.
29 Vijayaraghavan K, Crum S, Cherukuri S, *et al*. Association of impedance cardiography parameters with changes in functional and quality-of-life measures in patients with chronic heart failure. *Congest Heart Fail* 2004;10:22–7.