USING THE CLINICAL SUMMARY SCORE FROM THE KANSAS CITY CARDIOMYOPATHY QUESTIONNAIRE AS AN ENDPOINT IN CLINICAL TRIALS: PSYCHOMETRIC SUPPORT

Chad Gwaltney, PhD[1] , Brian Tiplady, PhD[1], Celine Deschaseaux, MPH[2]

[1] ERT, Pittsburgh, PA; [2] Novartis Pharma AG

ABSTRACT

Objectives

- Symptoms and physical limitations can have an important impact on the day to day lives of heart failure patients. The Clinical Summary Score (CSS) of the Kansas City Cardiomyopathy Questionnaire (KCCQ), a patient-reported outcome instrument, provides a measure of symptoms and physical limitations associated with heart failure. The primary goal of this study was to evaluate the psychometric properties of the KCCQ CSS and its utility as an endpoint in clinical trials.

Methods

- Data from 3 randomized, controlled clinical trials with heart failure patients were included in analysis (ALOFT, PARAMOUNT, and PARADIGM-HF trials). Studies were examined independently; within each study, data were collapsed across treatment groups. Study measures included the KCCQ, physician-rated New York Heart Association (NYHA) classification, patient global impression of change (PGIC), and/or NT-proBNP assay. RESULTS: Findings were similar across the 3 trials. Mean KCCQ CSS scores at baseline ranged from 63-76 (on a 0-100 scale, with higher scores indicating better symptoms and physical functioning). The CSS consistently discriminated between all four NYHA classifications (all pairwise comparisons p<.05). Correlations with BNP and NT-proBNP levels were statistically significant, but – as expected - relatively small (-.12 and -.17 respectively). The KCCQ CSS was sensitive to changes in patient status over time, as indexed by changes in the NYHA classification and PGIC. The responder definition – the amount of change within an individual patient that would be considered clinically meaningful - was in the range of 5 to 10 points, which is slightly higher than what has been seen in studies of the KCCQ overall score.

Conclusions

- Data from multiple heart failure clinical trials confirm the psychometric characteristics of the KCCQ CSS. This evidence supports the use of the KCCQ CSS as an endpoint in clinical trials examining heart failure treatments.

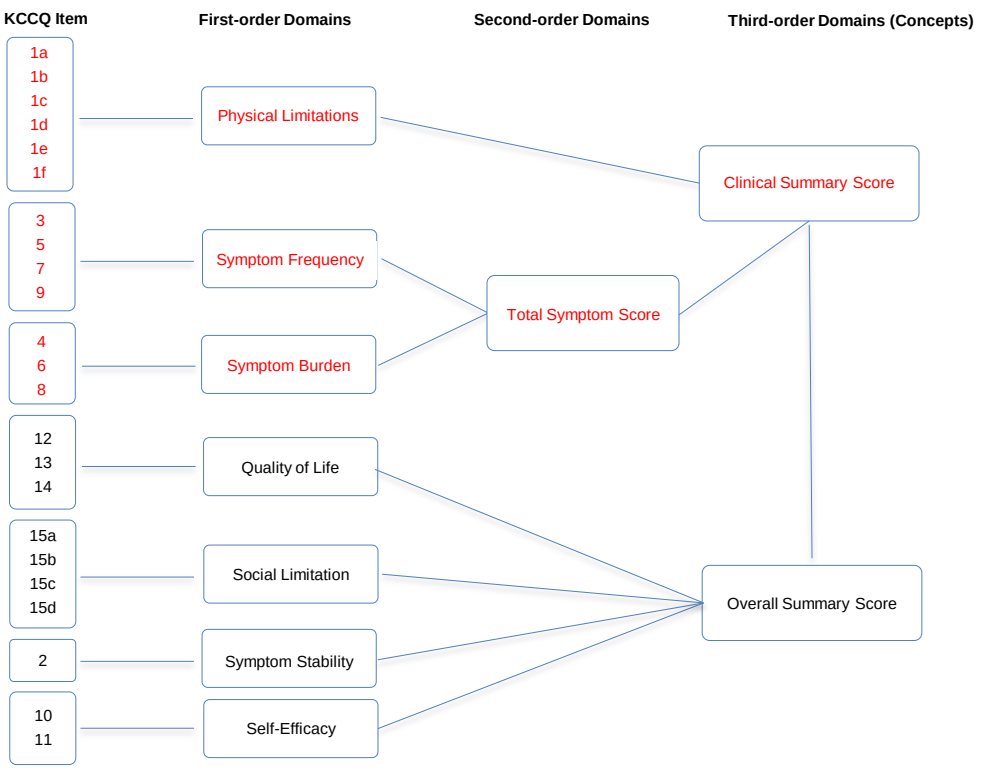
INTRODUCTION

- Heart failure is a complex clinical syndrome that results from any structural or functional impairment of ventricular filling or ejection of blood
- Heart failure patients experience a variety of debilitating symptoms – including dyspnea, fatigue, and peripheral edema - and physical limitations that disrupt their daily life and may be an indicator of poor prognosis (hospitalization, mortality)
- Heart failure (HF) may impact a patient's quality of life more than other chronic diseases, including diabetes and arthritis[1, 2]
- Reducing symptoms and physical limitations is an important goal for new HF therapies
- Patient-reported outcome instruments that adequately capture HF-related symptoms and physical limitations are needed to measure the effects of new therapies
- The Clinical Summary Score (CSS) of the Kansas City Cardiomyopathy Questionnaire (KCCQ) was developed to assess the most important symptoms and physical limitations associated with HF[3]
- The primary objective of this project was to assess the quantitative psychometric characteristics of the KCCQ CSS – including its reliability, validity, and sensitivity to change – and to estimate a responder definition that provides and index of clinically meaningful change
- These analyses may facilitate the use of the KCCQ CSS as an endpoint in HF clinical trials

KCCQ DESCRIPTION

- KCCQ: PRO instrument including 23 items that assess multiple dimensions of HRQOL
- KCCQ CSS : 13 items addressing total symptom and physical limitation domains
- KCCQ Physical Limitation Domain : 6 items addressing activities that can be affected by HF (e.g., dressing, walking, climbing stairs)
- KCCQ Total Symptom Domain: Includes 7 items addressing frequency and bother associated with HF-related symptoms (dyspnea, swelling, fatigue)
- Scale of 0-100 with higher scores indicating better health

Figure 1. Conceptual Framework of the KCCQ



CLINICAL TRIALS INCLUDED IN ANALYSIS

KCCQ CSS quantitative psychometrics assessed using data from 3 clinical trials:

1) ALOFT: A twelve-week, randomized, double-blind, multi-center, placebo controlled, parallel group study to evaluate the safety and efficacy of aliskiren 150 mg when added to standard therapy in patients with stable heart failure
2) PARAMOUNT: A 36-week, randomized, double-blind, multi-center, parallel group, active controlled study to evaluate the efficacy, safety and tolerability of LCZ696 compared to valsartan in patients with chronic heart failure and preserved left ventricular ejection fraction
3) PARADIGM-HF: A multicenter, randomized, double-blind, parallel group, active-controlled study to evaluate the efficacy and safety of LCZ696 compared to enalapril on morbidity and mortality in patients with chronic heart failure and reduced ejection fraction

Table 1. Patient Characteristics **(Patients with KCCQ Data)**

Patient Characteristic	ALOFT (N= 278)	PARAMOUNT (N= 308)	PARADIGM-HF (N= 7784)
Age (Mean/SD)	68.3 (10.2)	71.1 (9.1)	64.2 (11.2)
Gender	78% Male	56% Female	79% Male
Race	97% Caucasian 2% Black 1% Asian	74% Caucasian 13% Asian 11% Other 2% Black	72% Caucasian 13% Asian 8% Other 5% Black
Ethnicity	>90% Other	68% Other 17% Hisp/Latino 13% Indian	69% Other 18% Hisp/Latino 7% Indian
NYHA Class	60% Class II 40% Class III	1% Class I 80% Class II 19% Class III	4% Class I 70% Class II 25% Class III 1% Class IV

Table 2. Assessments and Timing of Administration*

Measure	ALOFT	PARAMOUNT	PARADIGM-HF
KCCQ	BL and 3 months	BL, 3, and 9 months	BL, 4, 8, 12, 24 and 36 months
New York Heart Association (NYHA) Functional Classification	BL and 3 months	BL, 3, and 9 months	BL, 4, 8, 12, 24 and 36 months
EQ-5D	Not administered	Not administered	BL, 4, 8, 12, 24 and 36 months
Patient Global Impression of Change (PGIC)	Not administered	BL, 3, and 9 months	4, 8, 12, 24 and 36 months
BNP/NT-proBNP	BL and 3 months	Not administered	Not administered

* Only includes administration times that were used in CSS psychometric analyses; BL = Baseline

REFERENCES

[1]Stewart AL, Greenfield S, Hays RD et al. Functional status and well being of patients with chronic conditions. Results from the Medical Outcomes Study. *JAMA* 1989;262:907-913.

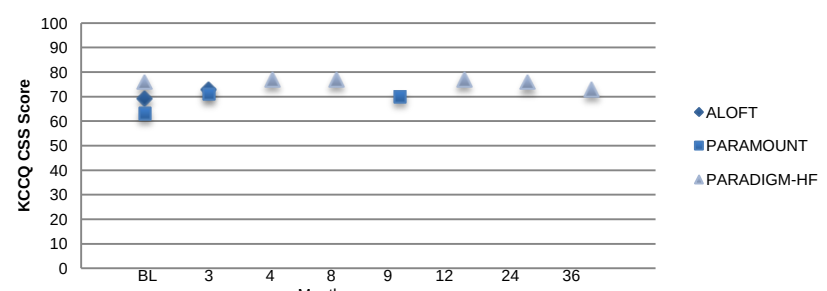
[2] Hobbs FDR, Kenkre JE, Roalfe AK, Davis RC, Hare R, Davies MK. Impact of heart failure and left ventricular systolic dysfunction on quality of life: A cross sectional study comparing common chronic cardiac and medical disorders and a representative adult population. *Eur Heart J* 2002;23:1867-1876.

[3] 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;35:1245-1255.

ANALYSIS PLAN

- Analyses examined key psychometric characteristics of the KCCQ CSS, in order to support this subdomain as an independent endpoint
- Data from each trial analyzed separately
- Data aggregated across treatment groups within each trial
- Reliability
  - Internal consistency of KCCQ CSS
  - Test-retest reliability (ICC) using stable patients on PGIC/NYHA
- Convergent/Discriminant Validity
  - Correlations with EQ-5D/BNP/NT-proBNP
- Known-Groups Validity
  - Association with NYHA (ANOVA)
- Sensitivity to Change
  - Association with change in NYHA/PGIC (effect size)
- Responder Definition
  - Average CSS score associated with improvement/worsening in NYHA/PGIC

Figure 2. Average KCCQ CSS Scores



Average KCCQ CSS scores at baseline and follow-up time points; Higher scores = better symptoms and functioning

Table 3. Test-Retest Reliability

Study*	CSS BL	CSS Time 2**	Difference	ICC***
PARAMOUNT (n= 74)	61.5	63.1	1.7	0.81
PARADIGM-HF (n=2036)	74.4	73.5	-0.9	0.80

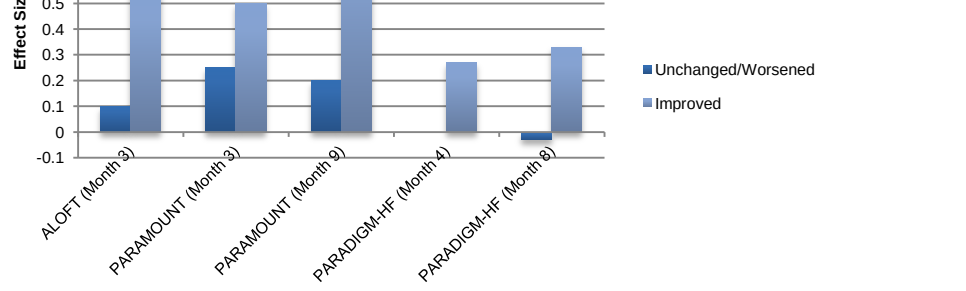
Conclusion: Examination of mean changes and ICCs supports reliability of KCCQ CSS (ICC ≥ 0.70 is common threshold for acceptable test-retest reliability); * Includes patients who were stable from BL to Time 2 on both the PGA and NYHA scales;
** Time 2 is Month 3 for PARAMOUNT and Month 4 for PARADIGM-HF; ***ICC = Intraclass Correlation Coefficient

Table 4. Internal Consistency

Study (Baseline data only)	Internal Consistency Estimate (Cronbach's alpha)
ALOFT	0.90
PARAMOUNT	0.92
PARADIGM-HF	0.92

Conclusion: Data support the internal consistency reliability of the KCCQ CSS (α ≥ 0.80 is common threshold for acceptable internal consistency)

Figure 3. KCCQ CSS Decreases As NYHA Classification Increases



* Lower KCCQ CSS scores indicate worse symptoms and physical limitations

Table 5. Relationship Between KCCQ CSS and EQ-5D Domains

EQ-5D Domain	Correlation (PARADIGM-HF only)*
Mobility	-0.54
Self-Care	-0.49
Usual Activities	-0.59
Pain/Discomfort	-0.44
Anxiety/Depression	-0.39
Visual Analogue Scale (Overall Health)	0.46

Conclusion: Increased impairment on EQ-5D subdomains associated with increased symptoms and physical limitations (lower scores) on KCCQ CSS; Greater overall health associated with less symptoms and physical limitations *All p<.01

Table 9. Sensitivity to Change in Patient Status

Biomarker	Correlation (ALOFT only)*
BNP	-0.12
NT-proBNP	-0.17

Conclusion: As expected, BNP/NT-proBNP increases are modestly associated with increased symptoms and physical limitations *p<.05

Figure 4. Sensitivity to Change in Patient Status

Conclusion: Changes in KCCQ CSS over time reflect self-reported global change ratings supporting the sensitivity of the measure
*Effect Size = Avg. Follow-Up CSS – Avg. BL CSS/Standard Deviation of BL Scores
*Improved/Unchanged/Worsened = Change on NYHA and/or PGA

Table 6. Responder Definition Analysis

Study (Time)	CSS FU – BL (Improved-No Change) *	CSS FU – BL (Worsened – No Change)*	1 SEM**	.5 SD
ALOFT	7.4	-4.1	6.3	9.9
PARAMOUNT (Month 3)	7.0	3.6	5.9	10.4
PARAMOUNT (Month 9)	5.8	-5.6		
PARADIGM-HF (Month 4)	1.1***	-8.9	8.7	9.7
PARADIGM-HF (Month 8)	1.3***	-10.4		

Conclusion: KCCQ CSS responder definition lies between 5-10 points
* FU = Follow-Up, BL = Baseline; Change based on PGA for PARAMOUNT and PARADIGM-HF and NYHA change in ALOFT
** SEM = Standard Error of Measurement
***Small estimates possibly due to relatively high KCCQ CSS scores at baseline

CONCLUSIONS

- The KCCQ CSS exhibited strong quantitative psychometric characteristics
- Internal consistency was high supporting the fidelity of the KCCQ CSS
  - Additional analyses to confirm the hierarchical structure of the scale (e.g., confirmatory factor analysis to examine first, second, and third order factors) may be useful
- The responder definition – the threshold for change within an individual patent that can be considered clinically meaningful appeared to lie within a small range (5-10 points)
  - Studies of the KCCQ overall score suggested a point estimate of 5 points
  - PARADIGM-HF estimates for patients who improved were lower than 5 points, but this may be due to the high BL scores observed in this trial
- KCCQ CSS may be a valuable endpoint in clinical trials
  - Captures important aspects of the patient's experience
  - Statistically efficient by capturing both symptoms and physical limitations in single score
  - Allows for more comprehensive benefit-risk assessment

ERT Innovating Better Health™

NOVARTIS