Canadian Journal of Cardiology 31 (2015) 792–799

Systematic Review/Meta-analysis

Baseline Functional Class and Therapeutic Efficacy of Common Heart Failure Interventions: A Systematic Review and Meta-analysis

Robert J.H. Miller, MD, Jonathan G. Howlett, MD, Derek V. Exner, MD, MPH, Patricia M. Campbell, MD, Andrew D.M. Grant, MD, and Stephen B. Wilton, MD, MSc

Libin Cardiovascular Institute of Alberta, University of Calgary, Calgary, Alberta, Canada

ABSTRACT

Background: New York Heart Association (NYHA) functional class provides important prognostic information and is often used to select patients for cardiovascular therapies, yet, the effect of NYHA class on therapeutic efficacy has not been systematically studied.

Methods: In this systematic review and meta-analysis we compared the relative and absolute mortality benefit of 5 common heart failure interventions (angiotensin-converting enzyme [ACE] inhibitors, β-blockers, mineralocorticoid receptor antagonists [MRAs], implantable cardioverter defibrillator [ICD], and cardiac resynchronization therapy [CRT]) across NYHA class. We included 26 randomized clinical trials of these interventions that reported all-cause mortality stratified according to baseline NYHA class in 36,406 patients.

Results: Pooled relative risk for NYHA I/II vs III/IV strata were similar for ACE inhibitors (0.90 vs 0.88), β-blockers (0.72 vs 0.79), MRA (0.79 vs 0.75), and CRT (0.80 vs 0.80), with all heterogeneity $P > 0.8$. Conversely, ICD efficacy was greater for class I/II (relative risk, 0.65 vs 0.86, heterogeneity $P = 0.02$). The pooled absolute risk difference was smaller for NYHA I/II vs III/IV with ACE inhibitors (−0.02 vs −0.06, $P = 0.12$), β-blockers (−0.02 vs −0.05, $P = 0.047$), MRA (−0.03 vs −0.11,

RÉSUMÉ

Introduction : La classification fonctionnelle de la New York Heart Association (NYHA) fournit des informations importantes pour évaluer le pronostic et est souvent utilisée pour sélectionner les patients en vue de traitements cardiovasculaires. Cependant, les répercussions de la classification de la NYHA sur l'efficacité thérapeutique n'ont pas fait l'objet d'études systématiques.

Méthodes : Au cours de la revue systématique et de la méta-analyse, nous avons comparé les avantages sur la mortalité relative et absolue de 5 interventions habituelles pour le traitement de l'insuffisance cardiaque (inhibiteurs de l'enzyme de conversion de l'angiotensine [ECA], β-bloqueurs, antagonistes du récepteur minéralocorticoïde [ARM], défibrillateur cardioverteur implantable [DCI] et thérapie de resynchronisation cardiaque [TRC]) de toutes les classes de la NYHA. Nous avons inclus 26 essais cliniques aléatoires concernant ces interventions qui rapportaient la mortalité toutes causes confondues stratifiée selon la classification initiale de la NYHA de 36 406 patients.

Résultats : Le risque relatif global de la strate I/II de la NYHA *vs* la strate III/IV était similaire pour les inhibiteurs de l'ECA (0,90 *vs* 0,88), β-bloqueurs (0,72 *vs* 0,79), ARM (0,79 *vs* 0,75) et TRC (0,80 *vs* 0,80),

The New York Heart Association (NYHA) functional classification was proposed in 1928 for quantifying heart failure disease severity and has since undergone several revisions.[1] NYHA functional class is a simple method of quantifying functional capacity based on history alone and has clear prognostic value for patients with heart failure.[2,3] It is widely used because of familiarity, but has questionable reproducibility and poor reliability.[4,5]

Baseline NYHA class is also often used as an inclusion criterion in clinical trials testing the efficacy of heart failure therapies. Initial trials of new interventions are often conducted in patients with more advanced NYHA class, because advanced NYHA class predicts a higher absolute risk of relevant clinical outcomes, and results in a smaller required sample. Clinical guidelines then often adopt these entry criteria in their recommendations to avoid extrapolation of study results beyond the populations in which they have been proven to be efficacious. This conservative approach implies that heart failure symptom severity is directly linked with therapeutic efficacy. However, the mechanisms of action and physiologic effects of many heart failure therapies are independent of baseline symptom severity. As an example, in a recent study no association between baseline NYHA class and the effects of mineralocorticoid receptor antagonists (MRAs) on ejection fraction or symptoms was found.[6] Thus, we hypothesized that for most evidence-based therapies for chronic systolic heart failure, baseline NYHA class is unrelated

Received for publication November 18, 2014. Accepted December 27, 2014.

Corresponding author: Dr Stephen B. Wilton, GE64 3280 Hospital Dr NW, Calgary, Alberta T2N 4Z6, Canada. Tel.: +1-403-210-7102; fax: +1-403-210-9180.

E-mail: sbwilton@ucalgary.ca

See page 798 for disclosure information.

http://dx.doi.org/10.1016/j.cjca.2014.12.031

$P = 0.001$), and CRT (-0.01 vs -0.04, $P = 0.036$), but was similar across NYHA class for the ICD (-0.07 vs -0.05; $P = 0.27$).
Conclusions: Relative mortality reductions with most interventions were independent of baseline NYHA class. However, ICD efficacy was greater with NYHA I/II vs III/IV limitation, and absolute benefit was greater with higher NYHA class. For interventions other than the ICD, there is little evidence supporting use of NYHA class as a rigid criterion for selecting heart failure therapies.

toute hétérogénéité $P > 0,8$. En contrepartie, l'efficacité du DCI était plus grande pour la classe I/II (risque relatif, 0,65 *vs* 0,86, hétérogénéité $P = 0,02$). La différence du risque absolu global était plus petite pour la classe I/II de la NYHA *vs* la classe III/IV avec les inhibiteurs de l'ECA ($-0,02$ *vs* $-0,06$, $P = 0,12$), les β-bloqueurs ($-0,02$ *vs* $-0,05$, $P = 0,047$), les ARM ($-0,03$ *vs* $-0,11$, $P = 0,001$) et la TRC ($-0,01$ *vs* $-0,04$, $P = 0,036$), mais était similaire dans toutes les classes de la NYHA avec le DCI ($-0,07$ *vs* $-0,05$; $P = 0,27$).
Conclusions : Les réductions relatives de la mortalité liée à la plupart des interventions étaient indépendantes de la classification initiale de la NYHA. Cependant, le DCI a montré une plus grande efficacité à la classe I/II de la NYHA *vs* la classe III/IV, et un avantage absolu plus grand avec les classes supérieures de la NYHA. En ce qui concerne les interventions autres que le DCI, peu de données probantes soutenant l'utilisation de la classification NYHA comme critère de sélection rigoureux des traitements de l'insuffisance cardiaque existent.

to therapeutic efficacy assessed according to relative risk (RR) reduction, although it might remain as a valid marker of absolute risk, and thus inform therapeutic effectiveness.

To test this hypothesis, we conducted this systematic review and meta-analysis to compare the RR reduction, absolute risk difference (RD), and number needed to treat (NNT), for the outcome of all-cause mortality, with guideline-endorsed pharmacological and nonpharmacological therapies in patients with chronic systolic heart failure, stratified according to lesser (NYHA I/II) vs greater (NYHA III/IV) baseline functional limitation.

Methods

Data sources

This review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement.[7] We searched Medline and EMBASE from 1966 to February 2013 to identify randomized clinical trials conducted in patients with chronic systolic heart failure (defined as left ventricular ejection fraction $< 45\%$), published in any language. We designed separate search strategies for each of 6 interventions: angiotensin-converting enzyme inhibitor (ACEi), angiotensin II receptor blocker (ARB), β-blockers, MRA, implantable cardioverter defibrillator (ICD), and cardiac resynchronization therapy (CRT). The complete Medline search strategy for ACEi is provided in Supplemental Table S1. The search of electronic databases was supplemented by an additional search of references from each included study and published systematic reviews.

Study selection

Two authors (R.J.H.M., S.B.W.) independently screened the titles and abstracts of all records. We selected articles for independent, duplicate full text review if they were original reports of randomized controlled clinical trials. On full text review, we included studies that reported the outcome of all-cause mortality stratified according to baseline NYHA class, over at least 6 months of follow-up. Trials conducted in patients with an acute or recent myocardial infarction were excluded to improve population homogeneity. For trials that met the other inclusion criteria, but that did not report mortality data stratified according to NYHA class, we attempted to obtain this information by contacting the study authors and by searching the US Food and Drug Administration Web site.[8] Disagreements regarding inclusion status were resolved by consensus.

Data extraction and synthesis

For each included trial, we extracted the study design, baseline clinical characteristics, sample size, and mortality data. All eligible ARB trials reported a composite primary outcome of mortality or cardiovascular hospitalization rather than all-cause mortality. We were unable to obtain isolated data on all-cause mortality for these 3 trials, so they were excluded from the main analysis, and instead were analyzed separately using their reported primary outcome.

We assessed internal validity of each included study using the Cochrane Collaboration risk of bias tool.[9] Because the stratification of functional class was inconsistent, we grouped results for patients with NYHA I/II vs NYHA III/IV.

We used fixed effects models with inverse variance weighting to perform stratified meta-analyses comparing the RR and absolute RD in patients with baseline NYHA class I/II vs III/IV for each therapy. Inverse variance weighting is required for calculation of heterogeneity between subgroups. We did not pool results across interventions, because background therapy in control groups evolved substantially over the study period. Heterogeneity between studies, and between NYHA strata were assessed using the Cochran Q and the I^2 statistic. For each intervention, we calculated stratum-specific NNT, normalized to 2 years of follow-up to improve comparability, from the pooled absolute RD. Analyses were conducted using Stata version 13 (Stata Corp, College Station, TX), and all statistical tests were 2-sided.

Results

In Figure 1, the results of the article search and selection process are outlined. We retrieved 10,305 unique references, of which 120 met screening eligibility criteria and were reviewed in detail. Of these, 91 trials were excluded for the reasons listed in Figure 1. A total of 26 trials published