Polygenic Score for Beta-Blocker Survival Benefit in European Ancestry Patients with Reduced Ejection Fraction Heart Failure

David E. Lanfear
Jasmine A. Luzum
Rui cong She
Hongsheng Gui
Mark P. Donahue

See next page for additional authors

Follow this and additional works at: https://scholarlycommons.henryford.com/cardiology_articles
Authors
David E. Lanfear, Jasmine A. Luzum, Ruicong She, Hongsheng Gui, Mark P. Donahue, Christopher M. O'Connor, Kirkwood F. Adams, Sandra Sanders-van Wijk, Nicole Zeld, Micha T. Maeder, Hani N. Sabbah, William E. Kraus, Hans-Peter Brunner-La Rocca, Jia Li, and Keoki L. Williams
Polygenic Score for Beta-Blocker Survival Benefit in European Ancestry

Patients with Reduced Ejection Fraction Heart Failure

Running Title: Lanfear et al; Beta-Blocker Polygenic Score for Heart Failure

David E. Lanfear, MD, MS1,2; Jasmine A. Luzum, PharmD, PhD1,3; Ruicong She, MS1,4; Hongsheng Gui, PhD1; Mark P. Donahue, MD5; Christopher M. O’Connor, MD6; Kirkwood F. Adams, MD7; Sandra Sanders-van Wijk, MD, PhD8; Nicole Zeld, BS1; Micha T. Maeder, MD, PhD9; Hani N. Sabbah, PhD2; William E. Kraus, MD5; Hans-Peter Brunner-LaRocca, MD8; Jia Li, PhD1,4; L. Keoki Williams, MD, MPH1

1Center for Individualized and Genomic Medicine Research (CIGMA), Department of Internal Medicine, Henry Ford Health System, Detroit, MI; 2Heart and Vascular Institute, Henry Ford Health System, Detroit, MI; 3Department of Clinical Pharmacy, University of Michigan College of Pharmacy, Ann Arbor, MI; 4Department of Public Health Sciences, Henry Ford Health System, Detroit, MI; 5Division of Cardiology, Duke University, Durham, NC; 6Inova Heart and Vascular Institute, Falls Church, VA; 7Division of Cardiology, University of North Carolina, Chapel Hill, NC; 8Department of Cardiology, Maastricht University, Maastricht, Netherlands; 9Cardiology Department, Kantonsspital St. Gallen, St. Gallen, Switzerland

Address for Correspondence:
David E. Lanfear, MD, MS
Henry Ford Health System
2799 W Grand Blvd, K14
Detroit, MI 48202
Fax: 313-916-8799
Telephone: 313-916-6375
Email: dlanfe1@hfhs.org

This manuscript was sent to Kenneth B. Margulies, MD, Senior Guest Editor, for review by expert referees, editorial decision, and final disposition.
Abstract

**Background:** Beta-blockers (BB) are mainstay therapy for heart failure with reduced ejection fraction (HFrEF). However, individual patient responses to BB vary, which may be partially due to genetic variation. The goal of this study was to derive and validate the first polygenic response predictor (PRP) for BB survival benefit in HFrEF patients.

**Methods:** Derivation and validation analyses were performed in n=1,436 total HF patients of European descent and with EF <50%. The PRP was derived in a random subset of the Henry Ford Pharmacogenomic Registry (HFPGR; n=248), and then validated in a meta-analysis of the remaining patients from HFPGR (n=247), the TIME-CHF (n=431), and HF-ACTION trial (n=510). The PRP was constructed from a genome-wide analysis of BB*genotype interaction predicting time to all-cause mortality, adjusted for MAGGIC score, genotype, level of BB exposure, and BB propensity score.

**Results:** Five-fold cross-validation summaries out to 1000 SNPs identified optimal prediction with a 44 SNP score and cutoff at the 30th percentile. In validation testing (n=1188) greater BB exposure was associated with reduced all-cause mortality in patients with low-PRP score (n=251; HR=0.19 [95% CI=0.04-0.51], p=0.0075), but not high-PRP score (n=937; HR=0.84 [95% CI=0.53-1.3], p=0.448), a difference that was statistically significant (p interaction =0.0235). Results were consistent regardless of atrial fibrillation, EF (≤40% vs. 41-50%), or when examining cardiovascular death.

**Conclusions:** Among patients of European ancestry with HFrEF, a PRP distinguished patients who derived substantial survival benefit from BB exposure from a larger group that did not. Additional work is needed prospectively test clinical utility and to develop PRPs for other population groups and other medications.

**Key Words:** polygenic score; pharmacogenomics; beta blockers; precision medicine
WHAT IS NEW

- We generated and validated the first polygenic predictor of beta blocker survival benefit in HFrEF patients derived from unbiased, genome-wide genotyping data. Additional validation and then testing in clinical trials is needed. The current data is limited to European ancestry patients.

WHAT ARE THE CLINICAL IMPLICATIONS

- This genetic profile, once proven in trials, could be used to target beta blockers to the subset of HFrEF patients that are likely to benefit and allow a large number of patients to forego unnecessary treatment, fundamentally changing our treatment approach from beta blockers for all to genetically targeted therapy. Parallel work in other ancestral groups is urgently needed.

- The general research approach followed here is theoretically applicable to any disease and medication, which could help accelerate broad progress towards clinically useful precision medicine.
Nonstandard Abbreviations and Acronyms

SNP: Single Nucleotide Polymorphism
BB: Beta blocker
MAGGIC: Meta-analysis Global Group in Chronic Heart Failure
HFPGR: Henry Ford Pharmacogenomic Registry
HF-ACTION: Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training
TIME-CHF: Trial of Intensified vs Standard Medical Therapy in Elderly Patients With Congestive Heart Failure
GWAS: Genome wide association study
PRP: Polygenic Response Predictor
PRS: Polygenic Risk Score
The landmark beta-blocker (BB) clinical trials for heart failure with reduced ejection fraction (HFrEF) showed a significant reduction in the risk of death overall,\textsuperscript{1-3} but it is important to point out that individual patient responses to BB widely vary. For example, in one randomized clinical trial, the 95\% confidence interval for change in left ventricular ejection fraction (LVEF) with carvedilol was -11.1 to +32.9.\textsuperscript{4} Despite this well-recognized variability BB response,\textsuperscript{5} current guidelines for HFrEF correctly adopt a “one size fits all” approach, whereby all patients are recommended to be treated with the same target doses of BB,\textsuperscript{6} because clinical characteristics largely did not impact HFrEF patients’ response to BB therapy in terms of survival.\textsuperscript{1-3} A relatively recent possible exception to this is atrial fibrillation, which some studies have reported negates BB effectiveness in HFrEF.\textsuperscript{7, 8} Since clinical factors are largely unable to identify BB responders vs. non-responders, understandably much research has focused on other factors, such as genetic variation, to support precision medicine approaches for BB treatment decisions.\textsuperscript{9-11} These pharmacogenetic studies, mostly employing the candidate gene approach,\textsuperscript{12-16} generally support the concept that genetic variation impacts BB response in HFrEF patients but results are inconsistent and no clinically actionable markers are proven to date.\textsuperscript{17} Although there are some examples where one or two genetic variants (usually in the setting of altered drug metabolism) have sufficient impact on drug response to generate a clinical action (e.g. CYP2C19 and clopidogrel)\textsuperscript{18}, it now appears that common complex phenotypes are likely under the influence of many genetic loci, each variant acting with relatively small impact,\textsuperscript{19-21} and drug response may be similarly complex and polygenic in nature.

Polygenic risk scores have emerged as a method to aggregate the small effects of numerous genetic variants into a score that reflects the overall genetic risk of a phenotype of interest. For common complex traits, this appears to often capture enough variation for clinical
utility, where a smaller number of genome-wide significant “hits” have failed to do so.\textsuperscript{22, 23} Polygenic scores have now been developed for several common diseases, such as coronary disease and others,\textsuperscript{23, 24} and these will soon be explored for implementation of targeted population management. Despite this exciting new development, similar analytic methods have largely not yet been successfully adapted to drug response in the setting of prevalent disease. There are emerging examples published,\textsuperscript{10, 25-34} but to our knowledge, none applied to treatment of HFrEF. Limited adaptation of polygenic scores to drug-response may be due in part to methodologic challenges, such as lack of sufficiently detailed drug exposure data and complexities of analysis. Although cohorts with detailed drug exposure information tend to be smaller in size relative to recent GWAS, a polygenic score approach may offer enhanced power and an efficient approach for constructing polygenic drug-response scores could have broad impact on precision medicine and drug development. The objective of this study was to derive and validate the first PRP for BB-associated survival benefit in patients with HFrEF.

**Methods**

**Datasets and Overall Approach**

The current study was approved by the Henry Ford Hospital Institutional Review Board, and all participants provided written informed consent. Three datasets were used: the Henry Ford Heart Failure Pharmacogenomic registry (HFPGR);\textsuperscript{35} the Trial of Intensified vs Standard Medical Therapy in Elderly Patients With Congestive Heart Failure (TIME-CHF);\textsuperscript{36} and Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training (HF-ACTION).\textsuperscript{37} The data from HFPGR is publicly available via dbGaP (https://www.ncbi.nlm.nih.gov/gap). The remaining data that support the findings of this study may be made available upon reasonable request from
qualified researchers trained in human subject confidentiality protocols by contacting Dr. David E. Lanfear (dlanfe1@hfhs.org) at Henry Ford Hospital.

The patient flow from parent study to final analytic cohorts is shown in the consort diagram (Supplemental Figure 1.) Only self-reported white patients were included in this analysis. The parent studies are described briefly below and in detail elsewhere, and each had institutional review board/ethics approval. A diagram illustrating overall study design is presented in Figure 1. We first created a derivation group upon which the PRP would be constructed, followed by testing in multiple independent datasets to test its performance. Since the HFPGR had the most granular drug-exposure data (derived from pharmacy claims and time-updating), it was randomly divided into two halves: one for derivation (n = 248) and the other to be included in the validation (n = 247). The PRP was derived solely in the derivation subset of HFPGR. Thus, the validation data included the remaining half of the HFPGR (n = 247), TIME-CHF (n = 431), and HF-ACTION (n = 510) for a total of 1188 validation patients.

BB exposure was computed using two different methods depending on the cohort. For both HFPGR cohorts BB exposure was calculated from pharmacy claims (i.e. drug actually dispensed to patient) and was updated over time, as previously described, and briefly summarized in the Supplemental Methods section. For TIME-CHF and HF-ACTION BB exposure was calculated from the specific drug and dose at baseline, using the same dose-equivalence scheme as above but implemented without any updates over time and without information on medication dispensing (i.e. assumes patients were receiving the dose prescribed).

The HFPGR enrolled patients from October 2007 through March 2015 at Henry Ford Health System in Detroit, MI, USA. The overall goal of the HFPGR is to discover novel ways to better predict prognosis and response to HFrEF treatments. Patients aged 18 years or older were
included if they had health insurance coverage, met the definition for heart failure as defined by
the Framingham Heart Study,\textsuperscript{39} and had at least one documented left ventricular ejection fraction
(LVEF); a total of 1760 patients were enrolled. For the current analysis those with LVEF < 50%
and self-identified European ancestry were included (n=495). BB exposure was calculated from
pharmacy claims and was updated over time, as previously described\textsuperscript{35, 38}

TIME-CHF was a randomized controlled trial that took place at 15 centers in Switzerland
and Germany from January 2003 to June 2008 that enrolled a total of 622 patients. The primary
objective of the trial was to compare N-terminal pro-B-type natriuretic peptide (NTproBNP)–
guided versus symptom-guided treatment for heart failure. To be included in the trial, individuals
had to meet the following criteria: age > 60 years, diagnosed with systolic heart failure, New
York Heart Association (NYHA) class of II or greater, hospitalization for heart failure within the
year prior to enrollment, and an NTproBNP level 2 times the upper limit of normal. For the
current analysis, only patients consenting to participate in the genetic substudy with analyzable
data and baseline LVEF < 50% were included (n=431). BB exposure was calculated from the
specific drug and dose at baseline, with dose standardization performed as previously
described.\textsuperscript{35, 38} Heart failure therapy guided by NTproBNP did not significantly affect the
primary outcome.

HF-ACTION was a National Institute of Health funded randomized clinical trial that took
place at 82 centers within the United States, Canada, and France from April 2003 through
February 2007, enrolling a total of 2331 patients. The primary objective of the trial was to
compare usual HFrEF care to usual HFrEF care plus exercise training, but it included a genetic
sub-study. The trial included adult HFrEF patients with LVEF ≤ 35% and NYHA class II to IV
symptoms despite optimal heart failure therapy for at least 6 weeks. BB exposure was calculated
from the specific drug and dose at baseline, with dose standardization performed as previously described.\textsuperscript{35} In primary analysis, exercise training resulted in non-significant reductions in the primary end point of all-cause mortality or hospitalizations although in the pre-specified adjusted analysis, the exercise intervention significantly reduced the composite primary outcome. We felt this did not justify the need to adjust for trial arm assignment when modeling all-cause mortality (the primary outcome for the current study). Only participants of the genetic sub-study with quality genetic array data, complete clinical data and self-identified as European ancestry were included in the current analysis (n=510).

**Genotyping**

Blood samples were collected, processed and stored at -70°C. DNA was extracted using standard methods.\textsuperscript{40} Patients from all three studies were genotyped with the Axiom® Biobank array (Affymetrix, Santa Clara, CA). This array was designed to optimize genome-wide imputation, and it includes \( \sim 600,000 \) genetic variants with the following characteristics: \( \sim 300,000 \) genome-wide variants with minor allele frequencies \( >1\% \), \( \sim 250,000 \) low frequency \( (<1\%) \) coding variants from global exome sequencing projects, and an additional \( \sim 50,000 \) variants for improved African ancestry coverage. All genotyping was performed at the University of Michigan genotyping core lab with standard quality checks performed. In brief, single nucleotide polymorphisms (SNPs) with a minor allele frequency \( <0.05 \), those not in Hardy–Weinberg equilibrium (HWE \( p < 10^{-8} \)), multi-allelic sites, and ambiguous SNPs were deleted from the analysis. Samples with sex inconsistencies (i.e., between patient report and genetically determined) or duplicate genotyping were removed. All datasets underwent imputation using the University of Michigan imputation server with Minimac\textsuperscript{341} and using the cosmopolitan reference panel from the 1000 Genomes Project.\textsuperscript{42} Imputed SNPs passing a quality threshold of \( r^2 > 0.5 \) were included for analysis.
Polygenic Response Predictor (PRP) Construction

A combined SNP selection and multivariable data analysis strategy was formulated to derive the PRP based on a GWAS of the HFPGR-derivation cohort, summarized in Figure 2. The primary outcome was time to all-cause mortality. This was modeled using Cox regression focusing in on the SNP-by-BB exposure multiplicative interaction term (SNP*BB) and including the main effects of BB exposure and SNP as covariates. An additional two covariates were also included in the model; to adjust for established clinical risk factors for death we include the MAGGIC score (calculated without BB as a contributing variable) and to adjust for potential confounding by indication or treatment bias we included a BB propensity score. Thus the overall model was

\[ \text{Death} = \text{MAGGIC} + \text{BBpropensity} + \text{BBexposure} + \text{SNP} + \text{SNP*BBexposure} \]

The Q-Q plot of the GWAS result is shown in supplemental Figure 2. BB propensity score was calculated using logistic regression of baseline characteristic variables with the resulting output separated into quartiles and used as an ordinal covariate. BB exposure was implemented as a time-dependent covariate (this is specific to HFPGR). The SNPs were then ranked according to the p-value of SNP*BBexposure and the selected SNPs were pruned down to the most significant one within each linkage disequilibrium (LD) block, which was kept for subsequent analysis.

The PRP is a linear combination of SNPs, in which the SNP weights are determined by the estimated Cox regression coefficients in the original model (from the derivation cohort). To select the optimal number of SNPs to include, a series of scores were created using Cox models similar to the above (including MAGGIC, Propensity, BBexposure, SNP and SNP*BBexposure) but sequentially adding more and more SNP and SNP*BBexposure terms, starting with the most highly associated locus and then incrementally adding one SNP at a time, sorted according to the p-value. The SNPs were coded with an additive genetic model (0, 1, or 2 for minor allele count).
These scores were tested using 5-fold cross-validation with the merge method to select the optimal number of SNPs that would maximize the time-dependent AUC for 1-year survival (estimated from the Cox model). The optimal set of SNPs (and the associated coefficients) were then used to build the final PRP.

To prepare for validation, the PRP was also dichotomized into “responder” and “non-responder” groups. The purpose of converting the continuous PRP into dichotomous categories was to facilitate clinical testing and future implementation of the PRP (i.e., to be able to identify sub-groups of patients where a specific therapeutic action is clearly defined). To accomplish this, we examined the BB survival benefit across a variety of score cutoff thresholds separating “responder” vs. “non-responders”, the performance of each tested in the HFPGR derivation cohort. Performance was judged by examining the BB HRs for high and low PRP patients (separately) with the cutoff at increasing deciles of the PRP within the derivation cohort (i.e. trying the 10th percentile as the cutoff for low vs high, then 20th percentile, and so on). The HR in each PRP group and their separation was examined. Separation was best at 30th percentile in the derivation cohort and the raw score at this level was then carried forward throughout validation testing. For HF-ACTION, 3 SNPS of the PRP did not pass quality metrics, so these were excluded and the PRP was tabulated using the remaining 41 SNPs. The dichotomization point was recalculated using a matching (without those 3 SNPs) PRP in derivation set to identify the corresponding raw score, and this was used as the dichotomization point for HF-ACTION validation testing.

**Statistical Analysis and PRP Validation**

In validation, the association of BB exposure was tested in PRP stratified groups in each of the 3 validation datasets using Cox models of survival (time to all-cause mortality). Our primary
analysis consisted of models adjusted for MAGGIC (again without BB variable)\textsuperscript{43} and BB propensity score,\textsuperscript{44, 45} with BB exposure treated as a continuous variable. Given the relatively small sample sizes of the individual datasets, the results from each of the individual validation datasets were combined by meta-analysis using a fixed effect model to provide an overall validation of the PRP (n=1188). We also tested for heterogeneity within the studies of the meta-analysis using Cohran’s Q test and report this along with the I\textsuperscript{2} statistic. We then tested models inclusive of both PRP categories with an interaction term (PRP*BB) to calculate an interaction p value. We also present Kaplan-Maier curves in PRP stratified groups comparing baseline BB dose dichotomized into high vs. low exposure at a dose equivalent of <0.5 vs. \geq 0.5, corresponding to half the target dose of BB in pivotal clinical trials (e.g. total daily dosage of 100 mg for metoprolol or 25 mg for carvedilol).

Since BB was not randomly assigned, we included propensity score adjustment for baseline BB use in all analyses. The propensity score was constructed using logistic regression predicting baseline BB using the following variables: Age, sex, creatinine, ischemic etiology, stroke, COPD, peripheral vascular disease, atrial fibrillation, and hypertension. The propensity scores were grouped into quartiles and this was used as a covariate (0, 1, 2, 3) in all analytic models. To verify the effectiveness of the propensity score for mitigating bias we performed a balance check by tabulating the weighted standardized difference for each covariate for each study. These ranged from 0.01 to 0.31 (Supplemental Table 1), implying reasonably balanced covariates\textsuperscript{47}.

Secondary analyses included similar models stratified by history of atrial fibrillation, and (separately) by EF category (<40\% vs 40-49\%) and examining cardiovascular (CV) death. Atrial fibrillation and EF categories were examined in stratified Cox models with the same covariates.
as above. The time to CV death was analyzed using Fine-Gray models (to account competing risks) and including the same covariates as main survival analyses above. All BB exposures were constructed across agents using dose-equivalency as a proportion of the target dose of agent as previously described. For HFPG-R-test the BB exposure was time-updating measurements while for TIME-CHF and HF-ACTION the baseline BB dose was used.

Baseline characteristics are shown in total and across cohorts. Continuous variables were summarized by the mean and standard deviation and categorical variables were summarized by counts and percentages. They were compared among the datasets using analysis of variance (ANOVA) for continuous variables and Chi-square test for categorical variables. All statistical analyses were performed using R 3.6.1 (R Foundation, Vienna, Austria). P < 0.05 was considered statistically significant.

Results
Baseline characteristics of the 4 datasets (the 1 derivation and 3 validation, total n = 1,436) are summarized in Table 1. The cohorts significantly differed in relation to all variables assessed, except for the frequency of diabetes and stroke. Notably the median BB dose/exposure varied, with HF-ACTION having a greater median (0.56 dose equivalents, representing half of target dose) compared to the other validation sets (0.28, 0.25 respectively for HFPG-R and TIME-CHF). NTproBNP levels were greater in TIME-CHF and least in HF-ACTION. Average follow-up across datasets ranged from 785 to 985 days and there was a total of 332 deaths (23%).

The PRP was developed as described in the methods and it optimized after the inclusion of 44 genetic loci. The list of the 44 SNPs comprising the PRP (along with annotation) are provided in Supplemental Table 2. The interaction effect between each of them and BB exposure...
was associated with mortality ($p < 10^{-4}$). The percentile of PRP that resulted in the best separation between BB responders and non-responders in the HFPGR derivation dataset was the 30th percentile (raw score = 68.14) and this was used as the dichotomization point for validation testing. In HF-ACTION, 3 of the 44 SNPs of the PRP did not have usable genotype calls so the remaining 41 SNPs were used to tabulate the PRP for HF-ACTION participants and the 30th percentile (based on this 41-SNP score) was re-tabulated in the derivation set and used as the dichotomization point for HF-ACTION validation testing. Based on the results from the derivation dataset, patients with a low PRP were hypothesized to be the BB responders in the validation datasets, and patients with a high PRP were hypothesized to be BB non-responders.

The hazard ratios for BB exposure (as a continuous variable incorporating dose) in each of the validation datasets and in the overall validation group (combined via fixed-effect meta-analysis) are displayed in Figure 3. The corresponding survival curves (with BB exposure dichotomized at 50% target dose using the baseline dosing) are shown in Figure 4. The hazard ratio for BB exposure was numerically lower (more protective) in all the low-PRP groups than in the corresponding high-PRP groups, however, the confidence intervals were wide within the individual datasets, reaching significance in TIME-CHF. The entire validation group testing included 1,188 patients overall and a total of 279 deaths (23.5% mortality). This analysis revealed a very strong and statistically significant protective association for BB in the low-PRP group (HR = 0.19 [95% CI = 0.058 – 0.64] $p = 0.0075$), but no significant association in the high-PRP group (HR = 0.84 [95% CI 0.53 – 1.3] $p = 0.45$). Formal test of interaction indicated a significant difference between the BB effect between PRP groups (interaction $p = 0.0235$). Of note, PRP group (high vs. low) was not itself significantly associated with mortality (HR=0.83, $p=0.37$), and there was no difference in heartrate by PRP category within any of the cohorts (all
p>0.05). Tests for heterogeneity in the meta-analysis resulted in a Cochran Q of 2.18 (p=0.34) and I^2 statistic of 0.08, suggesting no significant heterogeneity.

We performed a series of additional sensitivity analyses. Since atrial fibrillation has been associated with impaired BB effectiveness in pooled subgroup analyses of BB pivotal trials, we performed additional modeling stratified by baseline history atrial fibrillation in the validation dataset. Overall, there were 364 patients with history of atrial fibrillation and 824 without. Among those without atrial fibrillation the low-PRP group showed a BB HR of 0.33 (p=0.11) while the corresponding high-PRP group BB HR was 0.73 (p=0.31). In HFrEF patients with comorbid atrial fibrillation, those with a low-PRP had a BB HR of 0.063 (p=0.017) while the comparable high-PRP group showed a BB HR of 0.96 (p=0.92). Similarly, since we included patients with EF<50% in our study but HFrEF has more recently been defined as EF <40%, we also performed secondary analyses stratified by EF category (<40% vs 40-49%). Overall, there were 971 patients with EF<40% and 217 with EF 40-49%, and in both subgroups the PRP trended towards predicting BB response with HR similar to the total validation analysis. In those with EF<40% and low-PRP the BB-HR was 0.23 (p=0.0203) while in the high-PRP patients BB-HR was 0.87 (p=0.57). In patients with baseline EF 40-49% and low-PRP the BB-HR was 0.053 (p=0.23) while in the corresponding high-PRP patients the BB-HR was 0.82 (p=0.76). Finally, we also tested PRP performance in terms of BB exposure association to cardiovascular (CV) death. There were a total of 200 CV death events modeled across the validation cohorts using a Fine-Gray model (to account competing risks) and the same covariates as above. The results are summarized in Table 2. In the total validation meta-analysis the HR of BB in low-PRP was 0.19 (95% CI 0.044 – 0.804) and HR of BB in the high-PRP was 0.89 (95% CI 0.51 – 1.5), and a test of interaction that was statistically significant (p=0.046).
Discussion

Neither previous candidate gene studies nor clinical predictors have been able to better target BB therapy from the broad indication achieved in HFrEF clinical trials. The PRP, a weighted calculation of genotypes at 44 loci across the genome, was able to do this in a validation group of more than 1100 patients. While polygenic scores specific for drug-response have been described, ours is the first validated example that we are aware of relevant to HF treatment, and one of the first in cardiovascular disease. Our results are provocative and require additional studies, but the potential implications are wide ranging.

Our approach and findings advance the existing literature on polygenic scores, which have been described for disease risk in variety of settings, including some cardiovascular diseases such as coronary artery disease and atrial fibrillation. Some investigators have taken these scores (for disease-risk) and then layered on drug therapy, while we have tried to go a step further by focusing on drug response (i.e. gene*drug interaction). This interaction approach has received less attention and presents unique challenges, such more complex statistical modeling and inherently lower power. The distinction is important however, because disease-risk (like pretest probability) is only one factor impacting treatment effectiveness, and existing evidence suggests distinct genetic architecture for drug response vs. disease-risk.

Approaches somewhat similar to ours have been described with varying levels of success in other disease areas, to which our data bring some additional optimism. Our sample size is not dissimilar from many early drug development programs, and the idea that greater definition of the “responder” subset could routinely be had for many medications is obviously very alluring. This approach could amplify effect sizes (shrinking sample size
requirements in pivotal trials), and ultimately reduce health care costs and adverse events by avoiding treatment in patients unlikely to be benefitted.

An inspection of our results raises a few observations that are worthy of note. One is that our BB PRP appeared of modest effect in HF-ACTION while the effect estimates in the other two validation cohorts were more dramatic, with the PRP being most statistically significant in TIME-CHF. The individual validation sets each had small sample sizes, so substantial variability across them is expected. Differences in the studies may also play a role. Specifically, HFPGR and TIME-CHF had more patients not treated with BB at baseline, and a lower average BB exposure, compared to HF-ACTION. In HF-ACTION, roughly 95% of patients in the parent study were treated with BB and indeed 94% of patients in the ACTION-HF subcohort of the current study. This reflects care more consistent with current guidelines, but the reduced variability in BB exposure (i.e. near uniform use) reduces power to detect gene-drug interactions which is impacted by how many patients are unexposed to treatment. On the other hand, the patients in the TIME-CHF were the highest risk of the 3 studies and it included the highest number and proportion of deaths, providing relatively greater power to detect a differences in BB-associated survival compared to the other two datasets. Despite all this, since each of the individual datasets is likely underpowered the more important consideration is that the overall validation result was quite clear and the direction of the association was consistent across all 3 datasets. Another point of potential interest is that the cut-off for favorable BB-response optimized at roughly 30% of the population. This suggests that most of the benefit of BB occurs in roughly one third of the population while the majority of patients show little, if any, benefit. This may seem counterintuitive for a therapy with population average benefit, but mathematically a smaller ‘super-responder’ group is a plausible explanation for an overall
benefit, and proportion is consistent with some data showing that a marked and sustained EF response to BB occurs in only a minority of patients. The prospect of identifying a smaller, hyper-responsive subgroup for BB (or eventually other HF treatments) via polygenic scores is tempting to speculate about since this could allow individualized drug therapy and reduce the typical polypharmacy of HF; many patients do not tolerate target doses of all indicated agents and these data highlight a potential solution to this real-world difficulty.

Since the PRP was derived by unbiased methodology (not based on pre-existing molecular biological knowledge) it may not be surprising that the loci of interest do not reflect established BB associations; a phenomenon previous seen in pharmacogenomic GWAS. Given the number of loci that inform the PRP and the fact that some are of unknown genomic function, a full interrogation of the mechanisms potentially at play remains a future endeavor, however, some pathways that indicate potential biologic plausibility are worth briefly noting. For instance, the top genes from pathway analysis of the PRP include ABCC3, a transporter that is associated with multidrug resistance, has been shown to alter propranolol cellular efflux, and could theoretically impact BB absorption or transport. The second gene of interest, ATP2B2a, generated two loci of the PRP and is a plasma membrane Ca\textsuperscript{+2} transporter. While a specific mechanism is not established, it is well known that Ca homeostasis is critical in heart failure and is impacted by BB therapy. Finally, variation in PDE5A was also a contributor to PRP and this gene’s product has well known cardiovascular effects though specifics to BB response in HF are unknown.

Our study has some limitations worthy of discussion. First, the BB PRP was only derived and validated in HFrEF patients of European ancestry. Thus, this BB PRP does not apply to patients of other ancestries. While we have previously shown that overall BB response appears
similar by race and by genomic ancestry, the individual genetic loci that determine this are likely to be very different across these groups due to different linkage patterns or even differing biologic drivers of response. A critical future undertaking is the construction of BB PRP for other ancestral groups (e.g. African or Asian ancestry), or ideally, a score robust to ancestry. While we provide the current functional annotations for the SNPs from bioinformatic databases, future experiments establishing the function and mechanisms of the SNPs and gene products would be needed in order to establish mechanism of the PRP. Another limitation is that the datasets utilized were observational regarding BB use. While BB randomized clinical trials would have been the ideal substrate for our work, this is not possible since it would have been unethical to randomize HFrEF patients to BB vs. placebo, and there is no available genome-wide data from the early pivotal BB trials. We attempted to minimize the treatment bias and other potential confounders by adjusting all models for baseline clinical risk (via MAGGIC risk score) and as well as BB propensity score. Lastly, while the relatively smaller size of the derivation set makes possible that our PRP or cut-off point could be further improved, we feel the strategy of keeping the completely distinct validation group relatively larger and including all 3 parent studies therein was preferred in order to maintain adequate validation sample size to test for survival differences and maximize the rigor and generalizability of our findings.

In summary, we developed and validated the first polygenic score for BB survival benefit in HFrEF patients of European ancestry. The score was able to differentiate patients with a greatly enhanced BB-associated survival benefit from a larger group of patients that did not feature a statistically significant survival benefit. These findings challenge the “one size fits all” approach for BB treatment in HFrEF and are a step toward precision medicine for HFrEF. Additional replication studies would be reassuring, and work to develop PRPs for other ancestral...
groups and other medications is warranted. Ultimately, prospective studies would be needed to establish the clinical utility of our BB PRP compared to the current standard of care, ideally via randomized trials.

**Funding Sources:** This research was supported by the National Heart, Lung, and Blood Institute (Lanfear R01HL103871, R01HL132154; Williams R01HL118267, R01HL141845; Sabbah P01HL074237, R01HL132154; Luzum K08HL146990, L30HL110279). Dr. Williams is also supported by the National Institute of Allergy and Infectious Diseases (R01AI079139) and the National Institute of Diabetes and Digestive and Kidney Diseases (R01DK064695, R01DK113003). Dr. Luzum was also supported by a Futures Grant from the American College of Clinical Pharmacy.

**Disclosures:** David E. Lanfear is a consultant for Amgen, Janssen, Ortho Diagnostics and DCRI (Novartis) and has participated in clinical trials from Amgen, Bayer, and Janssen, and has submitted a provisional patent request for the beta blocker polygenic score. Jasmine A. Luzum has nothing to disclose. Ruicong She has nothing to disclose. Hongsheng Gui has nothing to disclose. Nicole Zeld has nothing to disclose. Mark P. Donahue has nothing to disclose. Christopher M. O’Connor has nothing to disclose. Kirkwood F. Adams has nothing to disclose. Sandra Sanders-van Wijk has nothing to disclose. Micha T. Maeder has nothing to disclose. Hani N. Sabbah has nothing to disclose. William E. Kraus has nothing to disclose. Hans-Peter Brunner-LaRocca has nothing to disclose. Jia Li has nothing to disclose. L. Keoki Williams has nothing to disclose.
References


Bohnke M, Abecasis GR, Mohlke KL, Cupples LA. Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet*. 2009;41:56-65


10.1161/CIRCHEARTFAILURE.119.007012


44. Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41-55


Table 1. Baseline characteristics of the total, derivation and validation datasets.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall (n = 1,436)</th>
<th>HFPGR-Derivation (n=248)</th>
<th>HFPGR-Test (n=247)</th>
<th>HF-ACTION (n=510)</th>
<th>TIME-CHF (n=431)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>69 (12)</td>
<td>71 (11)</td>
<td>71 (10)</td>
<td>61 (12)</td>
<td>76 (7.6)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Female sex</td>
<td>397 (28%)</td>
<td>72 (29%)</td>
<td>77 (31%)</td>
<td>100 (20%)</td>
<td>148 (34%)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>MAGGIC (no BB)</td>
<td>21.9 (7.30)</td>
<td>18.8 (6.66)</td>
<td>17.9 (7.29)</td>
<td>20.9 (6.57)</td>
<td>27.2 (5.24)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>LVEF (%)</td>
<td>30.3 (9.7)</td>
<td>34.9 (10.7)</td>
<td>37.3 (9.6)</td>
<td>25.1 (7.1)</td>
<td>29.7 (7.7)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>28.8 (6.3)</td>
<td>30.6 (7.0)</td>
<td>30.8 (7.3)</td>
<td>29.8 (5.8)</td>
<td>25.4 (4.2)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Heart Rate</td>
<td>73 (13.7)</td>
<td>70 (12.3)</td>
<td>70 (12.9)</td>
<td>70 (11.3)</td>
<td>76 (14.3)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>SBP (mmHg)</td>
<td>120 (20)</td>
<td>126 (23)</td>
<td>126 (20)</td>
<td>115 (19)</td>
<td>118 (18)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>NTproBNP (ng/L)</td>
<td>3837 (5099)</td>
<td>3070 (3085)</td>
<td>3079 (2916)</td>
<td>1889 (2549)</td>
<td>6473 (7216)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Creatinine (mg/dL)</td>
<td>1.28 (0.62)</td>
<td>1.19 (0.63)</td>
<td>1.17 (0.57)</td>
<td>1.33 (0.75)</td>
<td>1.34 (0.44)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Ischemic</td>
<td>849 (59%)</td>
<td>135 (54%)</td>
<td>140 (57 %)</td>
<td>327 (64%)</td>
<td>247 (57%)</td>
<td>0.033</td>
</tr>
<tr>
<td>COPD</td>
<td>273 (19%)</td>
<td>56 (23%)</td>
<td>59 (24%)</td>
<td>71 (14%)</td>
<td>87 (20%)</td>
<td>0.002</td>
</tr>
<tr>
<td>Atrial fibrillation</td>
<td>450 (31%)</td>
<td>86 (35%)</td>
<td>98 (40%)</td>
<td>127 (25%)</td>
<td>139 (32%)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>History Hypertension</td>
<td>730 (79%)</td>
<td>215 (87%)</td>
<td>211 (85%)</td>
<td>NA</td>
<td>304 (70%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Stroke</td>
<td>104 (7.2%)</td>
<td>15 (6.1%)</td>
<td>22 (8.9%)</td>
<td>32 (6.3%)</td>
<td>35 (8.1%)</td>
<td>0.432</td>
</tr>
<tr>
<td>Peripheral Vascular</td>
<td>167 (11.6%)</td>
<td>14 (5.6%)</td>
<td>25 (10.1%)</td>
<td>41 (8%)</td>
<td>87 (20.2%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Diabetes</td>
<td>486 (34%)</td>
<td>94 (38%)</td>
<td>91 (37%)</td>
<td>154 (30%)</td>
<td>147 (34%)</td>
<td>0.119</td>
</tr>
<tr>
<td>Follow-up (days)</td>
<td>887 (499)</td>
<td>844 (642)</td>
<td>785 (519)</td>
<td>988 (380)</td>
<td>852 (499)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Any Beta Blocker</td>
<td>1143 (80%)</td>
<td>161 (65%)</td>
<td>157 (64%)</td>
<td>480 (94%)</td>
<td>345 (80%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Carvedilol</td>
<td>472 (41%)</td>
<td>66 (41%)</td>
<td>45 (29%)</td>
<td>279 (58%)</td>
<td>82 (24%)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Metoprolol Succinate</td>
<td>417 (36%)</td>
<td>57 (35%)</td>
<td>57 (36%)</td>
<td>134 (28%)</td>
<td>169 (49%)</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Metoprolol Tartrate</td>
<td>127 (11%)</td>
<td>34 (21%)</td>
<td>44 (28%)</td>
<td>49 (10%)</td>
<td>0 (0%)</td>
<td></td>
</tr>
<tr>
<td>Bisoprolol</td>
<td>67 (5.9%)</td>
<td>0 (0%)</td>
<td>0 (0%)</td>
<td>6 (1.3%)</td>
<td>61 (18%)</td>
<td></td>
</tr>
<tr>
<td>Other beta-blockers</td>
<td>60 (5.2%)</td>
<td>4 (2.5%)</td>
<td>11 (7%)</td>
<td>12 (2.5%)</td>
<td>33 (9.6%)</td>
<td></td>
</tr>
<tr>
<td>Proportion target dose</td>
<td>0.36 (0.37)</td>
<td>0.23 (0.29)</td>
<td>0.22 (0.28)</td>
<td>0.56 (0.43)</td>
<td>0.28 (0.27)</td>
<td>&lt; 0.001</td>
</tr>
<tr>
<td>Death</td>
<td>332 (23.1%)</td>
<td>53 (21.4%)</td>
<td>36 (14.6%)</td>
<td>78 (15.3%)</td>
<td>165 (38.3%)</td>
<td>&lt; 0.001</td>
</tr>
</tbody>
</table>

BMI = body mass index; COPD = chronic obstructive pulmonary disease; HF-ACTION = the Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training; HFPGR = Henry Ford Heart Failure Pharmacogenomic Registry; LVEF = left ventricular ejection fraction; NTproBNP = N-terminal pro b-type natriuretic peptide; SBP = systolic blood pressure; TIME-CHF = Trial of Intensified vs Standard Medical Therapy in Elderly Patients With Congestive Heart Failure randomized trial.
Table 2. Cox model results for cardiovascular death using competing outcomes approach.

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Group</th>
<th>N</th>
<th>HR</th>
<th>P</th>
<th>P (interaction)</th>
</tr>
</thead>
<tbody>
<tr>
<td>HFPGR-test</td>
<td>Low PRP</td>
<td>45</td>
<td>0.03</td>
<td>0.14</td>
<td>0.11</td>
</tr>
<tr>
<td>HFPGR-test</td>
<td>High PRP</td>
<td>202</td>
<td>1.62</td>
<td>0.57</td>
<td></td>
</tr>
<tr>
<td>TIME-CHF</td>
<td>Low PRP</td>
<td>88</td>
<td>0.28</td>
<td>0.24</td>
<td>0.34</td>
</tr>
<tr>
<td>TIME-CHF</td>
<td>High PRP</td>
<td>343</td>
<td>0.84</td>
<td>0.68</td>
<td></td>
</tr>
<tr>
<td>HF-ACTION</td>
<td>Low PRP</td>
<td>118</td>
<td>0.17</td>
<td>0.073</td>
<td>0.16</td>
</tr>
<tr>
<td>HF-ACTION</td>
<td>High PRP</td>
<td>392</td>
<td>0.74</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>Meta-Analysis</td>
<td>Low PRP</td>
<td>251</td>
<td>0.19</td>
<td>0.024</td>
<td>0.046</td>
</tr>
<tr>
<td>Meta-Analysis</td>
<td>High PRP</td>
<td>937</td>
<td>0.89</td>
<td>0.68</td>
<td></td>
</tr>
</tbody>
</table>
Figure Legends

**Figure 1.** Schematic illustrating the use of the three datasets for polygenic response predictor (PRP) derivation and validation. HF-ACTION = the Heart Failure: A Controlled Trial Investigating Outcomes of Exercise Training; HFPGR = Henry Ford Pharmacogenomic Registry; TIME-CHF = the Trial of Intensified vs Standard Medical Therapy in Elderly Patients With Congestive Heart Failure; PRP = polygenic response predictor.

**Figure 2.** Schematic and formulae for calculating the PRP. AUC = area under the curve; BB = Beta Blocker; LD = linkage disequilibrium; MAGGIC = Meta-Analysis Global Group in Chronic Heart Failure risk score; PS = propensity score; QC = quality control; ROC = receiver operating characteristic curve; SNP = single nucleotide polymorphism.

**Figure 3.** Forest plot of hazard ratios for BB exposure in each of the validation cohorts and for the total validation (meta-analysis). The optimal PRP score cutoff from the derivation dataset was tested in Cox proportional hazards models adjusted for MAGGIC and BB propensity score. Low and high PRP indicate values above or below threshold (30th percentile of the derivation set). BB = Beta Blocker; PRP = polygenic response predictor.

**Figure 4.** Kaplan Meier survival curves stratified by PRP (high vs. low) and BB exposure for each dataset. The left panels show patients with a low PRP and the right-sided panels show patients with a high PRP. The red curves are patients with high BB exposure (≥50% target...
dose), and blue curves are patients with low BB exposure. BB = Beta Blocker; PRP = polygenic response predictor.
HFPGR-derivation (n = 248)

~7M SNPs post-QC

Each SNP was tested for interaction with BB in Cox proportional hazards models:

\[ \lambda(t) = \lambda_0 e^{a_1 + MGIC + a_2 \cdot PS + a_3 \cdot study + b_1 \cdot BB + b_2 \cdot SNP + w \cdot SNP \cdot BB} \]

Top 10k SNPs were clustered into LD blocks and the most significant SNP within each LD block was kept.

The top 1000 SNPs were evaluated for the combination that maximized the AUC ROC.

The 44 SNPs that maximized the ROC AUC were combined into the PRP using the following formula:

\[ \text{score} = \sum_{j=1}^{X} w_j \cdot SNP_j \]

PRP score cutoffs were varied by 10th percentiles and tested to identify the optimal cutoff separating "responders" from "non-responders".

The optimal PRP score cutoff was carried forward for validation (Figure 1).