

Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies

Stuart J. Pocock^{1*}, Cono A. Ariti¹, John J.V. McMurray², Aldo Maggioni³, Lars Køber⁴, Iain B. Squire⁵, Karl Swedberg⁶, Joanna Dobson¹, Katrina K. Poppe⁷, Gillian A. Whalley⁷, and Rob N. Doughty⁷, on behalf of the Meta-Analysis Global Group in Chronic Heart Failure (MAGGIC)

¹Department of Medical Statistics, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK; ²Institute of Cardiovascular and Medical Sciences, University of Glasgow, Glasgow, UK; ³ANMCO Research Centre, Florence, Italy; ⁴Rigshospitalet—Copenhagen University Hospital, Copenhagen, Denmark; ⁵Department of Cardiovascular Sciences, The University of Leicester, Leicester, UK; ⁶Sahlgrenska University, Hospital/Östra, Göteborg, Sweden; and ⁷Department of Medicine, University of Auckland, Auckland, New Zealand

Received 22 May 2012; revised 3 August 2012; accepted 13 September 2012; online publish-ahead-of-print 24 October 2012

See page 1391 for the editorial comment on this article (doi:10.1093/eurheartj/ehs363)

Aims

Using a large international database from multiple cohort studies, the aim is to create a generalizable easily used risk score for mortality in patients with heart failure (HF).

Methods and results

The MAGGIC meta-analysis includes individual data on 39 372 patients with HF, both reduced and preserved left-ventricular ejection fraction (EF), from 30 cohort studies, six of which were clinical trials. 40.2% of patients died during a median follow-up of 2.5 years. Using multivariable piecewise Poisson regression methods with stepwise variable selection, a final model included 13 highly significant independent predictors of mortality in the following order of predictive strength: age, lower EF, NYHA class, serum creatinine, diabetes, not prescribed beta-blocker, lower systolic BP, lower body mass, time since diagnosis, current smoker, chronic obstructive pulmonary disease, male gender, and not prescribed ACE-inhibitor or angiotensin-receptor blockers. In preserved EF, age was more predictive and systolic BP was less predictive of mortality than in reduced EF. Conversion into an easy-to-use integer risk score identified a very marked gradient in risk, with 3-year mortality rates of 10 and 70% in the bottom quintile and top decile of risk, respectively.

Conclusion

In patients with HF of both reduced and preserved EF, the influences of readily available predictors of mortality can be quantified in an integer score accessible by an easy-to-use website www.heartfailurerisk.org. The score has the potential for widespread implementation in a clinical setting.

Keywords

Heart failure • Meta-analysis • Prognostic model • Mortality

Introduction

Heart failure (HF) is a major cause of death, but prognosis in individual patients is highly variable. Quantifying a patient's survival prospects based on their overall risk profile will help identify those patients in need of more intensive monitoring and therapy, and also help target appropriate populations for trials of new therapies.

There exist previous risk models for patients with HF.^{1–8} Each uses a single cohort of patients and hence their generalizability to other populations is questionable. Each model's

development is from a limited cohort size, compromising the ability to truly quantify the best risk prediction model. Also most models are restricted to patients with reduced left-ventricular ejection fraction (EF), thus excluding many HF patients with preserved EF.

The Meta-analysis Global Group in Chronic Heart Failure (MAGGIC) provides a comprehensive opportunity to develop a prognostic model in HF patients, both with reduced and preserved EF. We use readily available risk factors based on 39 372 patients from 30 studies to provide a user-friendly score that readily quantifies individual patient mortality risk.

* Corresponding author. Tel: +44 207 927 2413, Fax: +44 207 637 2853, Email: stuart.pocock@lshtm.ac.uk

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author 2012. For permissions please email: journals.permissions@oup.com

Methods

The MAGGIC program’s details are documented previously.⁹ Briefly, we have individual patient data from 31 cohort studies (six randomized clinical trials and 24 observational registries). Here one registry is excluded since it had only median 3-month follow-up. The remainder comprised 39 372 patients with a median follow-up of 2.5 years (interquartile range 1.0–3.9 years), during which 15 851 patients (40.2%) died. Thirty-one baseline variables were considered as potential predictors of mortality (Table 1).

The Coordinating Centre at the University of Auckland assembled the database for 29 studies. The London School of Hygiene and Tropical Medicine team added in the CHARM trial data. The online Appendix lists the MAGGIC investigators (Supplementary material online).

In 18 studies, a preference was for rounding the EF to the nearest 5%. In these studies, such rounded values were re-allocated within 2.5% either side using a uniform distribution.

Statistical methods

Poisson regression models were used to simultaneously relate baseline variables to the time to death from any cause, with study fitted as a random effect. Since mortality risk is higher early on, the underlying Poisson rate was set in three time bands: up to 3 months, 3–6 months, and over 6 months. Models were built using forward stepwise regression with inclusion criterion $P < 0.01$.

For binary and categorical variables, dummy variables were used. Quantitative variables were fitted as continuous measurements, unless there was a clear evidence of non-linearity, e.g. body mass index, EF, and creatinine. Also two highly significant statistical interactions were included in the main model: the impact of age and systolic blood pressure both depend on EF.

Each variable’s strength of contribution to predicting mortality was expressed as the z statistic. The larger the z the smaller the P-value, e.g.: z values 3.29, 3.89, 5.32, and 6.11 are associated with P-values 0.001, 0.0001, 0.0000001, and 0.000000001, respectively.

Missing values are handled by multiple imputations using chained equations.^{10,11} This method has three steps. First, for each variable with missing values, a regression equation is created. This model includes the outcome and follow-up time, in this case the Nelson–Aalen estimator (as recommended by White and Royston¹⁰), an indicator variable for each study and other model covariates. For continuous variables, this is a multivariable linear regression, for binary variables, a logistic regression, and for ordered categorical variables, an ordinal logistic regression. Once all such regression equations are defined, missing values are replaced by randomly chosen observed values of each variable in the first iteration. For subsequent iterations, missing values are replaced by a random draw from the distribution defined by the regression equations. This was repeated for 10 iterations, the final value being the chosen imputed value. This is similar to Gibbs sampling.¹²

This entire process was repeated 25 times, thus creating 25 imputed data sets. The next step was to estimate the model for each of these data sets. Finally, the model coefficients are averaged according to Rubin’s rule.¹³ This ensures that the estimated standard error of each averaged coefficient reflects both between and within imputation variances, giving valid inferences.

We converted the Poisson model predictor to an integer score, which is then directly related to an individual’s probability of dying within 3 years. A zero score represents a patient at lowest possible risk. Having grouped each variable into convenient intervals, the score increases by

Table 1 Descriptive statistics for baseline variables

	Alive (n = 23 521)		Died (n = 15 851)	
	Mean or %	SD	Mean or %	SD
Age (years)	64.3	11.8	71.9	10.9
Male, %	69.0		65.1	
Non-Caucasian, %	10.7		7.8	
Body mass index (kg/m ²)	27.5	5.1	26.0	5.0
Current smoker, %	34.2		29.0	
Ejection fraction, %	36.6	14.0	33.6	14.0
Systolic blood pressure (mmHg)	131.0	21.8	130.5	25.6
Diastolic blood pressure (mmHg)	77.7	12.1	75.5	13.5
Haemoglobin (g/L)	133.7	19.0	119.0	26.1
Heart failure duration ≥ 18 months, %	48.8		49.7	
NYHA class, %				
I	10.8		6.7	
II	53.8		37.1	
III	31.3		42.8	
IV	4.1		13.4	
Creatinine (μmol/L)	109.4	55.8	126.9	58.4
Sodium (mmol/L)	139.7	3.6	138.9	4.2
Medical history, %				
Diabetes	20.6		25.7	
Angina	40.3		38.6	
MI	45.6		43.6	
Atrial fibrillation	17.8		23.5	
Stroke	6.2		12.2	
COPD	5.7		17.0	
Hypertension	41.3		39.3	
Rales	22.3		41.7	
Ischaemic heart disease	52.9		51.8	
CABG	15.4		13.9	
PCI	11.7		7.9	
Branch bundle block	22.1		24.5	
Oedema	21.4		31.9	
Shortness of breath, %				
Resting	15.9		35.8	
Exercise	80.8		78.8	
Medications, %				
Beta-blocker	40.4		24.4	
ACE-I	68.0		60.5	
ARB	3.3		4.3	

NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease; ACE-I, angiotensin-converting enzyme inhibitor; ARB, angiotensin-receptor blockers; PCI, percutaneous coronary intervention; CABG, coronary artery bypass grafting.

an integer amount for each risk factor level above the lowest risk. Each integer is a rounding of the exact coefficient in the Poisson model, making log rate ratio 0.1 equivalent to 1 point.

The data were analysed using Stata version 12.1 statistical package.

Results

This report is based on 39 372 patients from 30 studies: six were randomized controlled trials (24 041 patients) and 24 were registries (15 331 patients). Supplementary material online *Table S1* describes each of the 30 studies. Overall, 15 851 (40.2%) patients died during a median follow-up of 2.5 years. The six largest studies (DIAMOND,¹⁴ DIG,¹⁵ CHARM,¹⁶ and ECHOS¹⁷ trials and IN-CHF¹⁸ and HOLA¹⁹ registries) contributed 75.8% of patients and also 75.8% of deaths.

There were 31 baseline variables available for inclusion in prognostic models. *Table 1* provides their descriptive statistics for patients still alive and patients who died during follow-up.

Using Poisson regression models for patient survival with forward stepwise variable selection, adjusting for study (random effect) and follow-up time (higher mortality rate in early follow-up), we identified 13 independent predictor variables (*Table 2*). All were highly significant $P < 0.002$, and most were overwhelmingly significant, i.e. $P < 0.0001$.

Table 3 lists the extent of missing data for these 13 variables. A multiple imputation algorithm (see Methods) was used to

overcome this problem. Consequently, all results are based on average estimates across 25 imputed data sets.

For continuous variables, potential non-linearity in the prediction of survival was explored, as were potential statistical interactions between predictors. Hence the associations of EF, body mass index, and serum creatinine with mortality risk were, respectively, confined to EF $< 40\%$, body mass index $< 30 \text{ kg/m}^2$, and serum creatinine $< 350 \mu\text{mol/L}$. The mortality association of increased age was more marked with higher EF, whereas the inverse association of systolic blood pressure with mortality became more marked with lower EF.

Figure 1 displays the independent impact of each predictor on mortality risk. The impact of age (which varies with EF) is particularly strong, and hence is shown on a different scale to the other plots.

From the risk coefficients given in *Table 2*, an integer score has been created (*Figure 2*). For each patient, the integer amounts contributed by the risk factor's values are added up to obtain a total integer score for that patient. The bell-shaped distribution of this integer risk score for all 39 372 patients is shown in *Figure 3*. The median is 23 points and the range is 0–52 points, with 95% of patients in the range of 8–36 points. The curve in *Figure 3* relates a patient's score to their probability of dying within 3

Table 2 Multivariable model predicting mortality in all 39 372 patients

Variable	Rate ratio	95% CI	Log rate ratio	Z	P-value
Age (per 10 years)	1.154	(1.092, 1.220)	0.143	5.08	<0.0001
Males	1.115	(1.073, 1.159)	0.109	5.58	<0.0001
BMI (per 1 kg/m ² increase up to 30 kg/m ²) ^a	0.965	(0.959, 0.972)	-0.035	-10.10	<0.0001
Current smoker	1.159	(1.109, 1.210)	0.147	6.65	<0.0001
SBP (per 10 mmHg increase)	0.882	(0.855, 0.910)	-0.126	-7.85	<0.0001
Diabetes	1.422	(1.365, 1.481)	0.352	16.85	<0.0001
NYHA					
I	0.788	(0.732, 0.848)	-0.239	-6.35	<0.0001
II	1.000				
III	1.410	(1.354, 1.467)	0.343	16.75	<0.0001
IV	1.684	(1.580, 1.796)	0.521	16.05	<0.0001
Ejection fraction (per 5% increase up to 40%) ^a	0.581	(0.539, 0.627)	-0.542	-14.03	<0.0001
COPD	1.228	(1.152, 1.310)	0.206	6.36	<0.0001
HF duration >18 months	1.188	(1.139, 1.240)	0.173	7.96	<0.0001
Creatinine (per 10 $\mu\text{mol/L}$ up to 350 $\mu\text{mol/L}$)	1.039	(1.035, 1.042)	0.038	19.82	<0.0001
Beta-blocker	0.760	(0.726, 0.796)	-0.274	-11.77	<0.0001
ACE-I/ARB	0.908	(0.856, 0.963)	-0.096	-3.26	0.002
Interaction of ejection fraction and age ^b	1.040	(1.031, 1.049)	0.039	9.05	<0.0001
Interaction of ejection fraction and SBP ^c	1.012	(1.008, 1.017)	0.012	5.13	<0.0001

BMI, body mass index; SBP, systolic blood pressure; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease; HF, heart failure; ACE-I, angiotensin-converting enzyme inhibitor; ARB, angiotensin-receptor blockers.

^aThe BMI variable has a linear trend up to 30 kg/m², while above 30 kg/m² the risk is constant. Similarly, for ejection fraction, the risk is constant above 40%, and for creatinine risk is constant above 350 $\mu\text{mol/L}$.

^bThe interaction between ejection fraction and age indicates an extra 4% increase in mortality for each simultaneous 10-year increase in age and 5% increase in ejection fraction on top of the risks of ejection fraction and age considered independently, i.e. the protective effect of increased ejection fraction function diminishes as a patient ages (*Figure 1*).

^cThe interaction between ejection fraction and SBP indicates an extra 1.2% increase in mortality for each simultaneous 10 mmHg increase in SBP and 5% increase in ejection fraction on top of the risks of ejection fraction and SBP considered independently, i.e. the protective effect of increased ejection fraction function diminishes as a patient's SBP increases (*Figure 1*).

Table 3 Extent of missing data

Model variable	Studies with no data		Studies with some data		Total patients missing data
	Studies	Missing patients	Studies	Missing patients	
Age	0	0	0	0	0
Gender	0	0	0	0	0
BMI	17	14 515	13	2686	17 201
Current smoker	6	9166	24	448	9614
SBP	9	12 016	21	276	12 292
Diabetes	1	348	29	341	689
NYHA class	5	2503	25	1128	3631
Ejection fraction	6	3279	24	3558	6837
COPD	10	16 788	20	253	17 041
HF duration	20	11 679	10	1066	12 745
Creatinine	5	2800	25	17 245	20 045
Beta-blocker	3	7890	27	709	8599
ACE-I/ARB	1	97	29	649	746

BMI, body mass index; SBP, systolic blood pressure; NYHA, New York Heart Association; COPD, chronic obstructive pulmonary disease; HF, heart failure; ACE-I, angiotensin-converting enzyme inhibitor; ARB, angiotensin-receptor blockers.

years. For instance, scores of 10, 20, 30, and 40 have 3-year probabilities 0.101, 0.256, 0.525, and 0.842, respectively. *Table 4* details the link between any integer score and the probabilities of dying within 1 year and 3 years.

Figure 4 shows mortality over 3 years for patients classified into six risk groups. Groups 1–4 comprise patients with scores 0–16, 17–20, 21–24, and 25–28, respectively, approximately the first four quintiles of risk. To give more detail at higher risk, groups 5 and 6 comprise patients with scores 29–32 and 33 or more, approximately the top two deciles of risk. The marked continuous separation of the six Kaplan–Meier curves is striking: the 3-year % dead in the bottom quintile and top decile is 10 and 70%, respectively.

Regarding model goodness-of-fit, *Figure 5* compares observed and model-predicted 3-year mortality risk across the six risk groups. In the bottom two groups, the observed mortality is slightly lower than that predicted by the model, but overall the marked gradient in risk is well captured by the integer score.

Tables 5 and *6* show two separate models for patients with reduced and preserved left-ventricular function (EF <40 and ≥40%, respectively). For most predictors, the strength of mortality association is similar in both subgroups. However, the impact of age is more marked and the impact of lower SBP is less marked in patients with preserved left-ventricular function, consistent with the interactions in the overall model.

In this meta-analysis of 30 cohort studies, we explored between-study heterogeneity in mortality prediction. From fitting separate models for each study, we observe a good consistency across studies re the relative importance of the predictors (data not shown). We have also repeated the model in *Table 2*, now fitting study as a fixed effect (rather than a random effect). This reveals substantial between-study differences in mortality risk not explained by predictors in our model. However, a comparison of

the seven randomized trials with the 23 patient registries reveals no significant difference in their mortality rates.

Discussion

This study identifies 13 independent predictors of mortality in HF. Although all have been previously identified, the model and risk score reported here are the most comprehensive and generalizable available in the literature. They are based on 39 372 patients from 30 studies with a median follow-up of 2.5 years, the largest available database of HF patients. Also, we include patients with both reduced and preserved EF, the latter being absent from most previous models of HF prognosis.

Given the wide variety of different studies included, with a global representation, the findings are inherently generalizable to a broad spectrum of current and future patients. Conversion of the risk model into a user-friendly integer score accessible by the website www.heartfailurerisk.org facilitates its use on a routine individual patient basis by busy clinicians and nurses.

All 13 predictors in the risk score should be routinely available, though provision will be made in the website for one or two variables to be unknown for an individual. Note, the ‘top five’ predictors age, EF, serum creatinine, New York Heart Association (NYHA) class, and diabetes are important to know. The inverse association of EF with mortality is well established, and as previously reported,⁹ in above 40% there appears no further trend in prognosis. We included serum creatinine rather than creatinine clearance or eGFR. The latter involve formulae that include age, which would artificially diminish the huge influence of age on prognosis.

We confirm the association of body mass index with mortality,²⁰ but with a cut-off of 30 kg/m², above which there appears no further trend. While others report heart rate as a significant predictor of mortality,²¹ we find that once the strong influence of

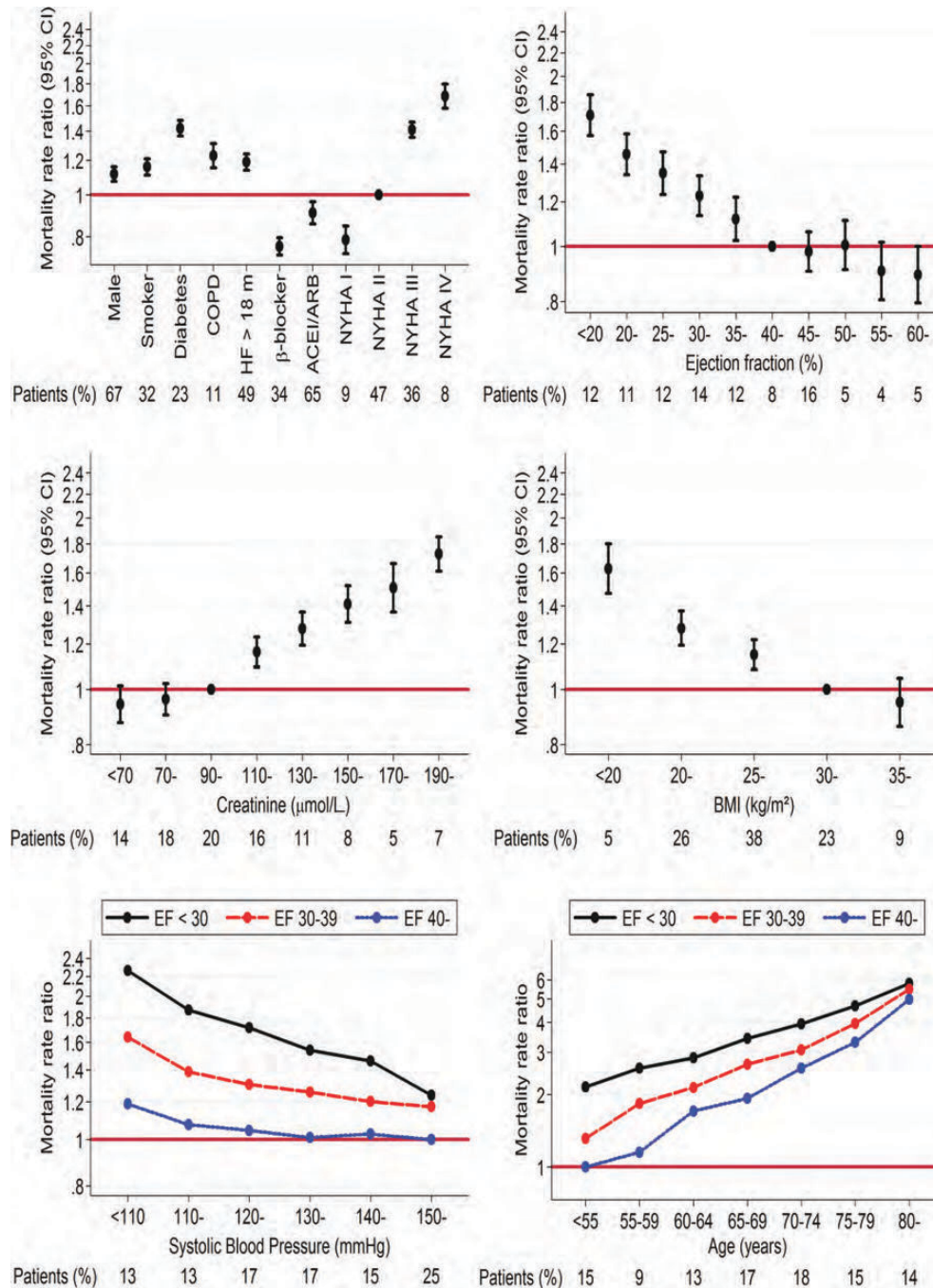


Figure 1 Mortality rate ratios (and 95% CIs) for each variable in the predictive model. All charts are on the same scale except that for the interaction between ejection fraction and age, where the impact on mortality is more marked.

beta blocker use is included, heart rate was not a strong independent predictor. A modest association of ACE-inhibitor and/or angiotensin-receptor blockers (ARB) use with lower mortality was highly significant, though many of our cohorts were established before ARBs were routinely available.

Cardiovascular disease history (e.g. myocardial infarction, angina, stroke, atrial fibrillation, LBBB) was considered in our model development. What mattered most was the time since

first diagnosis of HF, best captured by whether this exceeds 18 months. Besides the powerful influence of diabetes, the other disease indicator of a poorer prognosis was prevalence of COPD. Previous myocardial infarction, atrial fibrillation, and LBBB were not sufficiently strong independent predictors of risk to be included in our model.

For patients with reduced and preserved EF, we developed separate risk models (Tables 5 and 6). Nearly all predictors display a

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts



Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research



With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips



Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.