Validation of U.S. Mortality Prediction Models for Hospitalised Heart Failure in The United Kingdom and Japan

Short Title: Validation of Risk Models in Decompensated Failure

Toshiyuki Nagai, M.D.,[1, 2] Varun Sundaram, M.D.,[1, 2, 3,] Ahmad Shoaib, M.D.,[4]
Yasuyuki Shiraishi, M.D.,[5] Shun Kohsaka, M.D.,[5] Kieran J Rothnie, Ph.D.,[1]
Susan Piper, M.D.,[6] Theresa A McDonagh, M.D.,[6] Suzanna MC Hardman, Ph.D.,[7]
Ayumi Goda, M.D.,[8] Atsushi Mizuno, M.D.,[9] Mitsuaki Sawano, M.D.,[5]
Alan S Rigby, CStat,[10] Jennifer K Quint, M.D.,[1] Tsutomu Yoshikawa, M.D.,[11]
Andrew L Clark, M.D.,[4] Toshihisa Anzai, M.D.,[2, 12] and John GF Cleland, M.D.[4, 13]

1. National Heart & Lung Institute, Imperial College London, London, United Kingdom
2. Department of Cardiovascular Medicine, National Cerebral and Cardiovascular Center, Osaka, Japan
3. Harington Heart and Vascular Institute, University Hospitals Cleveland Medical Center, Case Western Reserve University, Cleveland, Ohio, USA & Royal Brompton and Harefield Hospitals, London, United Kingdom
4. Department of Cardiology, Hull York Medical School, Castle Hill Hospital, Kingston-upon-Hull, United Kingdom
5. Department of Cardiology, Keio University School of Medicine, Tokyo, Japan
6. Cardiology Department, King's College Hospital, London, United Kingdom
7. Clinical & Academic Department of Cardiovascular Medicine, Whittington Hospital, London, United Kingdom
8. Division of Cardiology, Kyorin University School of Medicine, Tokyo, Japan
9. Department of Cardiology, St, Luke's International Hospital, Tokyo, Japan
10. Department of Statistics, Hull York Medical School, University of Hull, Kingston-upon-Hull, United Kingdom
11. Department of Cardiology, Sakakibara Heart Institute, Tokyo, Japan
12. Department of Cardiovascular Medicine, Hokkaido University Graduate School of Medicine, Sapporo, Japan
13. Robertson Centre for Biostatistics & Clinical Trials, University of Glasgow and National Heart & Lung Institute, Royal Brompton & Harefield Hospitals, Imperial College London, London, United Kingdom

***Corresponding author*:**

Toshiyuki Nagai M.D., Ph.D., and John GF Cleland M.D., National Heart & Lung Institute, Imperial College London, London, United Kingdom, Emmanuel Kaye Building, Manresa Road, London, SW3 6LR, United Kingdom, TEL: +44(0)2075948821, E-mail: t.nagai@imperial.ac.uk; j.cleland@imperial.ac.uk

Conflict of Interest:

Toshiyuki Nagai; None, Varun Sundaram; None, Ahmad Shoaib; None, Yasuyuki Shiraishi; None, Shun Kohsaka; None, Keiran J Rothnie; None, Susan Piper; None, Prof Theresa McDonagh; None, Dr Suzanna Hardman; None, Ayumi Goda; None, Atsushi Mizuno; None, Mitsuaki Sawano; None, Prof Alan Rigby; None, Jennifer K Quint; None, Tsutomu Yoshikawa; None, Prof Andrew L Clark; None, Prof Toshihisa Anzai; None, Prof John GF Cleland; received grants and honoraria from Amgen, Novartis, Medtronic, Philips, Servier and Stealth Biotherapeutics.

Funding: Dr. Nagai is supported by grants from the Daiichi Sankyo Foundation of Life Science and the Mochida Memorial Foundation for Medical and Pharmaceutical Research. This research was supported by a Butterfield Award for UK-Japan collaboration in Medicine & Health from The Great Britain Sasakawa Foundation.

Abstract: (248 words; Limit 250 words)

Aims: Prognostic models for hospitalised heart failure (HHF) were developed predominantly for patients of European origin in the United States of America; it is unclear whether they perform similarly in other health-care systems or for different ethnicities. We sought to validate published prediction models for HHF in the United Kingdom (UK) & Japan.

Methods and Results: Patients in the UK (894) and Japan (3,158) were prospectively enrolled and similar in terms of sex (~60% men) and median age (~77 years). Models predicted that British patients would have a higher mortality than Japanese, which was indeed true both for in-hospital [4.8% vs 2.5%] and 180-day [20.7% vs 9.5%] mortality. The model c-statistics for the published/derivation [range 0.70-0.76] and Japanese [range 0.75-0.77] cohorts were similar and higher than for the UK [0.62-0.75] but models consistently over-estimated mortality in Japan. For in-hospital mortality, OPTIMIZE-HF performed best, providing similar discrimination in published/derivation, UK and Japanese cohorts [c-indices: 0.75 (0.74-0.77); 0.75 (0.68 - 0.81) and 0.77 (0.70 - 0.83)], and least over-estimated mortality in Japan. For 180-day mortality, the c-statistics for ASCEND-HF were similar in published/derivation [0.70] and UK [0.69 (0.64 - 0.74)] cohorts but higher in Japan [0.75 (0.71 - 0.79)]; calibration was good in the UK but again over-estimated mortality in Japan.

Conclusion: Calibration of published prediction models appear moderately accurate and unbiased when applied to British patients but consistently overestimate mortality in Japan. Identifying the reason why patients in Japan have a better than predicted prognosis is of great interest.

Keywords: acute heart failure; hospitalised heart failure; mortality prediction; outcome; Japan

Introduction

The prevalence of heart failure is rising due to an ageing population and longer survival after the onset of cardiovascular diseases, including heart failure itself.[1-4] Improvements in care have failed to stem a rising tide of heart failure related hospital admissions. Annually, heart failure is the primary reason for >200,000 admissions in Japan,[2, 5] >80,000 in the United Kingdom (UK)[6] and about one million in the United States of America (USA).[7] Despite advances in the management of chronic heart failure, mortality amongst patients hospitalized with worsening heart failure remains high and no intervention has been convincingly shown to improve outcome. [8-15]

Prognostic models for hospitalized heart failure (HHF) derived from surveys, registries, and randomized clinical trials have identified many variables that are associated with outcome. [16-27] Knowing a patient's risk may help guide management, including the intensity of follow-up, the urgency of advanced interventions or the need for palliative care. Moreover, some prognostic variables, such as renal function or serum potassium concentration, might be therapeutic targets. However, prognostic models have been developed primarily in patients of European origin and in the USA. It is unclear whether these published models predict outcome in other geographic regions with different health-care systems or other ethnic groups. Important regional differences exist not only in terms of health economy, medical infrastructure and management but also patient characteristics and their adherence to medical advice and lifestyle.[28-31] Accordingly, we investigated the validity of five published HHF mortality prediction models for cohorts of patients enrolled in the UK and Japan.

Methods

Data sources

The UK HHF Cohort comprised registry data from two National Health Service (NHS) hospitals in London and one in Kingston-upon-Hull,[32] each serving a local population of approximately 250,000 people and participating in the NHS England and Wales National Heart Failure Audit.[6] Between 2011 and 2013, 894 patients were prospectively enrolled. Data were generally acquired within hours of admission. This was part of a national survey initiated by the NHS which provided ethical oversight.

The WET-HF (WEst Tokyo Heart Failure) registry is an ongoing, multicenter, prospective observational registry of HHF in five large academic medical centers in metropolitan Tokyo (East Japan) that enrolled 2407 patients between 2011 and 2015.[33] The NaDEF (National cerebral and cardiovascular center for acute DEcompensated heart Failure) registry enrolled 751 HHF patients prospectively at a single centre in Osaka (West Japan) between 2013 and 2015 based on the same inclusion/exclusion criteria as those in the WET-HF registry.[34] Data in both registries were also generally acquired within hours of admission. The study protocols of the both the Japanese registries were approved by the respective institutional review boards, and were registered at the Japanese UMIN Clinical Trial Registration (UMIN000001171 and UMIN000017024, respectively).

Patients were enrolled shortly after admission and mortality was recorded from admission providing it is reasonably accurate.

Mortality prediction models and study population

We carried out a detailed search using the MEDLINE/PubMed and EMBASE search engines and identified all the HHF mortality prediction models based on a specific search strategy

(**Supplemental Appendix)**. This search strategy has been previously validated with high sensitivity and specificity for finding prediction research in MEDLINE.[35] Among these 28 models, five HHF models[18, 19, 21, 26, 27] were chosen based on the available covariates available in both the UK and Japanese datasets (15 models were excluded due to lack of specific variables which are required to calculate the risk score and 8 models were excluded due to lack of availability of in-hospital mortality and 180 day outcomes) (**Figure 1**) for external validation in the UK and Japanese cohorts (**Table 1 and Supplementary Tables S1-1, S2, S3-1 and S4-1**).

The HHF models identified were; 1) The Get With the Guidelines-Heart Failure (GWTG-HF) risk score,[18] 2) The Acute Decompensated Heart Failure National Registry (ADHERE) model,[21] 3) The Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF) risk score,[19] 4) The Acute Study of Clinical Effectiveness of Nesiritide in Decompensated Heart Failure (ASCEND-HF) risk score,[27] 5) The Outcomes of a Prospective Trial of Intravenous Milrinone for Exacerbations of Chronic Heart Failure (OPTIME-CHF) risk score.[26]

For each prediction model, we replicated the methods used by the original authors to calculate the predicted mortality for patients in the UK and Japanese cohorts. Racial origin was not collected but the vast majority of patients in the UK were of European origin and in Japan were Japanese. Accordingly, we did not use the ethnicity variable in the model, which applies mainly to people of African-American origin.

In the UK, blood urea nitrogen (BUN) was missing in 41% of cases, reducing the population available for analysis of all models other than OPTIME-CHF, which used serum creatinine. Therefore, we repeated all other models, substituting BUN with serum creatinine using the formula shown in **Supplemental Tables S1-2, S3-2 and S4-2**. In the ADHERE registry we

used the coefficients of BUN, systolic BP, heart rate (HR) and age from their multivariable model to calculate the log-odds of in-hospital mortality (0.0212×BUN-0.0192×Systolic BP+0.0131×HR+0.0288×age-4.72) as shown in the original paper. The OPTIME-CHF risk score was developed to predict 60-day mortality but because 60-day mortality was not recorded for the UK, the score was applied for 90-day mortality in each data-set.

41% (N=373) of UK patients had missing data on BUN and 27% (N=845) of Japanese patients did not have a record of New York Heart Association (NYHA) class. We therefore compared baseline characteristics and outcomes in British patients with and without BUN and in Japanese patients with and without a record of NYHA class. Survival status on days 30, 90 and 180 day was recorded rather than the precise date of death. Some patients did not have data on survival status on day 30 (N=68 in the UK, N=792 in Japan), on day 90 (N=29 in the UK, N=844 in Japan) or on day 180 (N=29 in the UK, N=906 in Japan) from the date of admission.

Statistical Analysis

Continuous variables are presented as mean ± standard deviation when normally distributed and otherwise as median and interquartile range (IQR). Comparison of differences between groups were made using the unpaired Student's t-test or Mann-Whitney U test for continuous variables, and using chi-squared test or Fisher's exact test for dichotomous variables, where appropriate. After defining predicted mortality (predicted probability of death) for each model using the risk scores or the log odds of mortality in the logistic regression model, we assessed discriminative performance and calibration for each model. Model discrimination was assessed by calculating the c-statistic with the area under the receiver operating characteristic curve. Calibration was assessed by plotting observed against predicted mortality, coefficient of

determination (large values of R^2 indicating a better fit) and the Hosmer-Lemeshow goodness of fit statistic (smaller p-values (eg:- P<0.01) indicating a poor fit). A high R^2 with a 'significant' p-value suggests that the prediction model has precision but lacks accuracy, which might be improved by introducing a 'correction factor'. All tests were two tailed, and a value of p <0.05 was considered statistically significant. All analyses were performed with Stata MP 14.2 (StataCorp., College Station, TX, USA).

Results

Cohort baseline characteristics

Baseline characteristics for British (N=894) and Japanese (N=3158) patients are summarised in **Table 2**. The median ages of British (76 [IQR 67-83] years) and Japanese (78 [IQR 69-84] years) patients were similar and approximately 60% of each cohort was men. Compared to Japanese patients, British patients were more likely to have severe HF as evidenced by worse NYHA classification, lower systolic BP, lower serum sodium concentrations and higher serum concentrations of BUN and creatinine but British patients had higher body mass index. Median systolic BP in derivation cohorts of the ASCEND-HF and the OPTIME-CHF were 123 and 120 mmHg, respectively, which were more similar to British compared to Japanese patients. Median serum creatinine concentrations of patients in the GWTG-HF, the ASCEND-HF and the OPTIME-CHF were also similar to the British cohort at 1.3, 1.2 and 1.4 mg/dL, respectively but lower amongst the Japanese (1.1mg/dL). Patients in the derivation cohorts and the UK were more likely to have a reduced LVEF than in Japan. The prevalence of ischemic heart disease was higher in each of the derivation cohorts (46-59%) and amongst British patients (48%) compared to the

Japanese (25%).

Compared to Japanese patients, British patients had shorter LOHS (11 days vs 15 days P <0.001); 11.1% of British and 23.8% of Japanese patients were hospitalised for >30 days. British patients had higher in-hospital (4.8% vs 2.5%, P=0.001), 30-day (5.7% vs 2.6%, P <0.001; 90-day (13.6% vs 6.0% P <0.001) and 180-day mortality (20.7% vs 9.5%, P <0.001) (**Table 3**). Despite, a higher proportion of patients with HFrEF in the UK being discharged on guideline-directed medical therapy, post-discharge mortality was significantly lower in Japan (**Table 3 and Supplementary Table S5**).

In-Patient Mortality

Risk scores for in-patient mortality for the GWTG-HF, modified-GWTG-HF and OPTIMIZE-HF and the log odds for mortality of the ADHERE and modified-ADHERE were all higher for British patients (**Table 3**) consistent with their worse actual prognosis. Model c-statistics were similar for the derivation and Japanese cohorts [18, 19, 21] but, apart from the OPTIMIZE-HF, performed less well in the UK (**Table 4**). In the UK, model discrimination improved when BUN was replaced with creatinine, possibly because this increased the number of patients available for analysis but for Japanese patients, replacing BUN with creatinine tended to reduce discrimination (**Table 4**). Overall, the OPTIMIZE-HF model, which uses creatinine rather than BUN, provided the best discrimination with remarkably similar c-statistics for the derivation [19], British and Japanese cohorts (0.753, 0.752 and 0.767 respectively).

Calibration coefficients for in-hospital mortality (**Figure 2A-C**) were also superior for the OPTIMIZE-HF compared to other models in both the UK (R^2 = 0.76) and Japan (R^2 = 0.91). Tests for goodness-of-fit consistently over-estimated mortality in all models for Japan, especially for

patients at lower risk. The OPTIMIZE-HF showed greater accuracy than other in-patient models for the UK and Japan. In Japan, substitution of BUN with creatinine reduced over-estimation of mortality leading to some improvement in fit with little effect on calibration coefficients (**Figure 3A and B**). Substitution of BUN with creatinine had little influence on model calibration in the UK cohort (**Figure 3A and B**).

Post-baseline Mortality at 30, 90 and 180 Days (**Table 4, Figure 3 D-F**).

For 30-day mortality (the ASCEND-HF), the c-statistic for the Japanese and derivation cohorts were similar[27] but lower in the UK. The model calibration coefficient was markedly higher for Japan (R^2 = 0.82) compared to the UK (R^2 = 0.12) but again over-estimated mortality in Japan. Substituting BUN with creatinine improved model precision in the UK but not Japan and improved calibration in the UK.

For 90-day (the OPTIME-CHF) mortality, the c-statistic for Japanese patients was again similar to that of the derivation cohort[26] but lower for British patients. Model coefficients were high for both cohorts but again over-estimated mortality in Japan; substituting BUN with creatinine had little effect on model performance for either cohort.

For 180-day mortality, the ASCEND-HF model c-statistics were similar for the derivation[27] and UK cohorts but higher for the Japanese. Model coefficients were high both for British (R^2 = 0.96) and Japanese (R^2 = 0.95) patients but again overestimated mortality in Japan. Substituting BUN with creatinine did not improve model performance.

Discussion

This analysis suggests that at least some published models for predicting mortality in HHF, developed predominantly for patients of European origin enrolled in the USA, provide similar prediction for patients enrolled in our registries in Japan and the UK. Indeed, the ability of the models to predict mortality was somewhat greater for Japanese compared to British patients. This suggests that the variables contributing to these models maintain their relationship to outcome in diverse health care systems and ethnicities. However, the published models consistently over-estimated mortality for HHF patients in Japan. These analyses should be confirmed and refined using other similar or larger data-sets internationally.

For in-patient mortality, the OPTIMIZE-HF performed well, in terms of discrimination and calibration, in both the UK and Japan. This is remarkable, given the very different lengths of hospital stay in the UK and Japan compared to the USA.[17, 19, 21, 36-40]. Longer hospital stays expose patients to a prolonged period at risk of events, even if they might subsequently lead to lower post-discharge mortality. Interestingly, the OPTIMIZE-HF was the only model that used serum creatinine in preference to BUN as a measure of renal dysfunction and adopted serum sodium as a "U-shaped" risk variable with 140 mEq/L as the nadir of risk.[19, 41] In the UK, improvement in the performance of in-patient and 30-day models by substituting BUN with creatinine might simply be attributed to the ability to include more patients in the analysis. In Japan, the effect of substituting BUN for creatinine in the GWTG-HF and the ADHERE in-patient models was complex; model discrimination declined slightly but model calibration improved somewhat for the GTWG-HF and over-estimation of mortality was reduced in both models. Model coefficients for the 90- and 180-day outcome models were high for both Japan and the UK but again consistently over-estimated mortality in Japan, which was a half to a quarter of that predicted. The high model

coefficients and consistently lower than predicted mortality for Japanese patients suggests that adding a variable to adjust for ethnicity might improve model calibration. In these longer-term models, substitution of BUN with creatinine had little effect on c-statistics or calibration and only modestly reduced over-estimation of mortality in Japan. The high model coefficients and consistently lower than predicted mortality for Japanese patients suggests that adding a variable to adjust for ethnicity might improve model calibration.

Other models for HHF suggest that urea is a better marker of prognosis than creatinine.[24] which is consistent with the reduction in model-discrimination after substituting BUN with creatinine in Japan. However, the prognostic superiority of BUN may be modest and creatinine may serve almost as well.[24] Indeed, serum urea, creatinine and their ratio may all provide additive prognostic information.[24, 42] The reason why models using BUN rather than serum creatinine should demonstrate lower model accuracy (discrimination and calibration) in British patients is uncertain. Serum creatinine reflects muscle creatinine turnover, protein intake and renal function.[43] BUN is also influenced by catabolic/anabolic balance but, in addition, a large proportion of urea is reabsorbed by the nephron especially if the patient is dehydrated or diuretic resistant. Greater muscle mass, differences in diet, the severity of congestion and use of higher doses of diuretics in the UK might all alter the relationships between BUN, creatinine and prognosis.

Health care systems and hospitalisation threshold vary widely across world regions; hence it is likely that risk scores provide different performance in different populations. In the present study, British patients had a worse risk profile compared to the Japanese, including lower average systolic BP, left ventricular ejection fraction and serum sodium concentration and worse renal function, although age was similar and BMI higher. The observed mortality for HHF in the UK

was similar to that predicted by the models. In other words, although the mortality associated with HHF may be higher in the UK than in some other countries, this reflects sicker patients, perhaps due to a higher threshold for admission. On the other hand, Japanese patients not only appeared less sick, perhaps reflecting a lower threshold for admission, but also had a much lower than predicted mortality, especially amongst lower risk patients. However, it is possible that models derived from "sicker" populations in the US overestimate mortality in "less sick" patients independent of the world region. We can only speculate as to the reasons for the observed difference in mortality, which may include differences in management, patients' adherence to medical advice and therapy, differences in culture, diet and lifestyle, heart failure aetiology, and genetics. Of note, mortality was higher in the UK despite a higher proportion of patients with HFrEF being discharged on guideline directed medical therapy (GDMT). One or more of the above differences might explain this anomaly.

Adding a variable to reflect a greater variety of ethnic groups might improve the performance of existing models.[39, 44] Alternatively, the influence of individual covariates could be reassessed in different ethnic groups and a new score developed but this might not capture the wider impact of ethnicity.

Prognostic models that can be applied internationally would have considerable value not only for research but also for auditing the quality of care. First, transfer of patients to a tertiary centre with advanced care / heart failure therapies such as left ventricular assisted device and transplantation can be done based on existing models of accurate mortality prediction. Second, development of effective community services for heart failure, a major thrust in the UK, means that many patients who used to be admitted to hospital for care are now managed at home or in day-care facilities. Accordingly, patients admitted to hospital are often sicker and more likely to

die. The National Heart Failure Audit for England & Wales reported an in-patient mortality for 32,991 HHF patients of 9.4% in 2013/14 (although < 3.0% if aged <65 years) which is a much higher mortality than reported in the USA (3.1% in 2010/2013)[45] or Japan (6.4% in 2007/2011).[28] Until an international model to assess prognosis is developed it may be impossible to determine whether poor outcome for HHF in a country reflects a high or low quality of care. Rather than starting afresh, it seems better to build on existing models, some of which performed well in this analysis. Starting with a model and only adding variables that improve it substantially or simplify it is likely to lead to faster and more certain evolution than re-inventing a new model for each new patient cohort.[24]

Limitations

Our study population includes patients only from selected academic medical centres in Japan and Britain which may not be representative for the whole population. The mortality in our study populations was lower than that observed in the published national data in both countries. However, the difference in mortality between the two countries in our study is quite similar to difference in mortality observed in national data-sets (~40% higher mortality in UK compared to Japan). Validating prediction models in national data-sets is not currently possible since these do not record required variables including blood pressure and laboratory data.

We used the same inclusion/exclusion HHF criteria for both Japanese registries (WET-HF and NaDEF) to increase the power of the analysis and generalizability of the results, although some differences in characteristics and outcomes were observed between these data-sets (**Supplementary Table S6**).

Renal function and NYHA functional class are key determinants of prognosis in patients

with HF.[46] The prediction models included in this analysis, with the exception of the OPTIMIZE-HF, used BUN as a measure of renal dysfunction but this variable was missing in many patients from the UK. Moreover, NYHA functional class was often missing in Japanese patients. Some significant differences in baseline characteristics and outcomes were observed amongst patients who did or did not have these data recorded. Patients who had a missing BUN appeared sicker while patients without a record of NYHA class appeared less sick (**Supplementary Table S7**). However, supplementary analysis restricted only to patients without missing BUN and NYHA data and using BUN showed similar results to the main analysis (**Supplementary Table S8 and Supplementary Figure S1*)***.

Fortunately, serum creatinine was available in >99% allowing many more patients to be included in the analysis when used instead of BUN. This may explain why the OPTIMIZE-HF model performed well in the UK and why substitution of BUN with creatinine improved the performance of some models for the UK cohort. The failure of similar substitution to improve discrimination in Japanese patients supports the notion that the improvement in the c-statistic in the UK was due to the inclusion of more patients rather than creatinine being a superior prognostic marker.

We excluded 23 published models from 28 identified mortality prediction models either because our data lacked specific variables that the models required or because they did not report on in-hospital or 180-day mortality.

The OPTIME-CHF model was developed for 60-day mortality, but as only 90-day mortality was recorded in the UK, we used this time-frame instead. As most deaths occurring in the first 90-days will have occurred within 60-days, this is unlikely to make a major difference.

The OPTIME-CHF and the ASCEND-HF models were developed on patients who

consented to participate in a randomized controlled trial. Such patients may not be epidemiologically representative and the effects of excluding patients who were not approached to participate or who declined is uncertain.

Conclusions:

Our analysis shows that existing prediction models predominantly derived from USA population and healthcare system provide fairly good discrimination for mortality amongst patients with HHF in the UK and Japan but overestimate mortality in Japan. Existing models could provide the basis for a universal mortality prediction model for HHF but might require modification depending on ethnicity. Further external validation of prediction models in diverse health care systems should be considered prior to application in routine clinical practice. Identifying the reason why patients in Japan have a better than predicted prognosis would be of great interest.

References

1. Roger VL, Weston SA, Redfield MM, Hellermann-Homan JP, Killian J, Yawn BP, Jacobsen SJ. Trends in heart failure incidence and survival in a community-based population. JAMA 2004;**292**(3):344-50.

2. Okura Y, Ramadan MM, Ohno Y, Mitsuma W, Tanaka K, Ito M, Suzuki K, Tanabe N, Kodama M, Aizawa Y. Impending epidemic: future projection of heart failure in Japan to the year 2055. Circ J 2008;**72**(3):489-91.

3. Levy D, Kenchaiah S, Larson MG, Benjamin EJ, Kupka MJ, Ho KK, Murabito JM, Vasan RS. Long-term trends in the incidence of and survival with heart failure. N Engl J Med 2002;**347**(18):1397-402.

4. He J, Gu D, Wu X, Reynolds K, Duan X, Yao C, Wang J, Chen CS, Chen J, Wildman RP, Klag MJ, Whelton PK. Major causes of death among men and women in China. N Engl J Med 2005;**353**(11):1124-34.

5. Yasuda S, Nakao K, Nishimura K, Miyamoto Y, Sumita Y, Shishido T, Anzai T, Tsutsui H, Ito H, Komuro I, Saito Y, Ogawa H, on the behalf of JI. The Current Status of Cardiovascular Medicine in Japan- Analysis of a Large Number of Health Records From a Nationwide Claim-Based Database, JROAD-DPC. Circ J 2016;**80**(11):2327-2335.

6. Cleland JG, McDonagh T, Rigby AS, Yassin A, Whittaker T, Dargie HJ, National Heart Failure Audit Team for E, Wales. The national heart failure audit for England and Wales 2008-2009. Heart 2011;**97**(11):876-86.

7. Munir MB, Sharbaugh MS, Thoma FW, Nisar MU, Kamran AS, Althouse AD, Saba S. Trends in hospitalization for congestive heart failure, 1996-2009. Clin Cardiol 2017;**40**(2):109-119.

8. Hsich EM, Grau-Sepulveda MV, Hernandez AF, Eapen ZJ, Xian Y, Schwamm LH, Bhatt DL, Fonarow GC. Relationship between sex, ejection fraction, and B-type natriuretic peptide levels in patients hospitalized with heart failure and associations with inhospital outcomes: findings from the Get With The Guideline-Heart Failure Registry. Am Heart J 2013;**166**(6):1063-1071 e3.

9. O'Connor CM, Starling RC, Hernandez AF, Armstrong PW, Dickstein K, Hasselblad V, Heizer GM, Komajda M, Massie BM, McMurray JJ, Nieminen MS, Reist CJ, Rouleau JL, Swedberg K, Adams KF, Jr., Anker SD, Atar D, Battler A, Botero R, Bohidar NR, Butler J, Clausell N, Corbalan R, Costanzo MR, Dahlstrom U, Deckelbaum LI, Diaz R, Dunlap ME, Ezekowitz JA, Feldman D, Felker GM, Fonarow GC, Gennevois D, Gottlieb SS, Hill JA, Hollander JE, Howlett JG, Hudson MP, Kociol RD, Krum H, Laucevicius A, Levy WC, Mendez GF, Metra M, Mittal S, Oh BH, Pereira NL, Ponikowski P, Tang WH, Tanomsup S, Teerlink JR, Triposkiadis F, Troughton RW, Voors AA, Whellan DJ, Zannad F, Califf RM. Effect of nesiritide in patients with acute decompensated heart failure. N Engl J Med 2011;**365**(1):32-43.

10. Packer M, O'Connor C, McMurray JJV, Wittes J, Abraham WT, Anker SD, Dickstein K, Filippatos G, Holcomb R, Krum H, Maggioni AP, Mebazaa A, Peacock WF, Petrie MC, Ponikowski P, Ruschitzka F, van Veldhuisen DJ, Kowarski LS, Schactman M, Holzmeister J, Investigators T-A. Effect of Ularitide on Cardiovascular Mortality in Acute Heart Failure. N Engl J Med 2017;**376**(20):1956-1964.

11. Massie BM, O'Connor CM, Metra M, Ponikowski P, Teerlink JR, Cotter G, Weatherley BD, Cleland JG, Givertz MM, Voors A, DeLucca P, Mansoor GA, Salerno CM, Bloomfield DM, Dittrich HC, Investigators P, Committees. Rolofylline, an adenosine A1-receptor antagonist, in acute heart failure. N Engl J Med 2010;**363**(15):1419-28.

12. Teerlink JR, Voors AA, Ponikowski P, Pang PS, Greenberg BH, Filippatos G, Felker GM, Davison BA, Cotter G, Gimpelewicz C, Boer-Martins L, Wernsing M, Hua TA, Severin T, Metra M. Serelaxin in addition to standard therapy in acute heart failure: rationale and design of the RELAX-AHF-2 study. Eur J Heart Fail 2017;**19**(6):800-809.

13. McMurray JJ, Teerlink JR, Cotter G, Bourge RC, Cleland JG, Jondeau G, Krum H, Metra M, O'Connor CM, Parker JD, Torre-Amione G, van Veldhuisen DJ, Lewsey J, Frey A, Rainisio M, Kobrin I, Investigators V. Effects of tezosentan on symptoms and clinical outcomes in patients with acute heart failure: the VERITAS randomized controlled trials. JAMA 2007;**298**(17):2009-19.

14. Teerlink JR, Cotter G, Davison BA, Felker GM, Filippatos G, Greenberg BH, Ponikowski P, Unemori E, Voors AA, Adams KF, Jr., Dorobantu MI, Grinfeld LR, Jondeau G, Marmor A, Masip J, Pang PS, Werdan K, Teichman SL, Trapani A, Bush CA, Saini R, Schumacher C, Severin TM, Metra M, Investigators REiAHF. Serelaxin, recombinant human relaxin-2, for treatment of acute heart failure (RELAX-AHF): a randomised, placebo-controlled trial. Lancet 2013;**381**(9860):29-39.

15. Mebazaa A, Nieminen MS, Packer M, Cohen-Solal A, Kleber FX, Pocock SJ, Thakkar R, Padley RJ, Poder P, Kivikko M, Investigators S. Levosimendan vs dobutamine for patients with acute decompensated heart failure: the SURVIVE Randomized Trial. JAMA 2007;**297**(17):1883-91.

16. Pocock SJ, Ariti CA, McMurray JJ, Maggioni A, Kober L, Squire IB, Swedberg K, Dobson J, Poppe KK, Whalley GA, Doughty RN, Meta-Analysis Global Group in Chronic Heart F. Predicting survival in heart failure: a risk score based on 39 372 patients from 30 studies. Eur Heart J 2013;**34**(19):1404-13.

17. Chioncel O, Mebazaa A, Harjola VP, Coats AJ, Piepoli MF, Crespo-Leiro MG, Laroche C, Seferovic PM, Anker SD, Ferrari R, Ruschitzka F, Fernandez SL, Miani D, Filippatos G, Maggioni AP, Investigators ESCHFL-TR. Clinical phenotypes and outcome of patients hospitalized for acute heart failure: the ESC Heart Failure Long-Term Registry. Eur J Heart Fail 2017.

18. Peterson PN, Rumsfeld JS, Liang L, Albert NM, Hernandez AF, Peterson ED, Fonarow GC, Masoudi FA, American Heart Association Get With the Guidelines-Heart Failure P. A

validated risk score for in-hospital mortality in patients with heart failure from the American Heart Association get with the guidelines program. Circ Cardiovasc Qual Outcomes 2010;**3**(1):25-32.

19. Abraham WT, Fonarow GC, Albert NM, Stough WG, Gheorghiade M, Greenberg BH, O'Connor CM, Sun JL, Yancy CW, Young JB, Investigators O-H, Coordinators. Predictors of in-hospital mortality in patients hospitalized for heart failure: insights from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). J Am Coll Cardiol 2008;**52**(5):347-56.

20. O'Connor CM, Abraham WT, Albert NM, Clare R, Gattis Stough W, Gheorghiade M, Greenberg BH, Yancy CW, Young JB, Fonarow GC. Predictors of mortality after discharge in patients hospitalized with heart failure: an analysis from the Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure (OPTIMIZE-HF). Am Heart J 2008;**156**(4):662-73.

21. Fonarow GC, Adams KF, Jr., Abraham WT, Yancy CW, Boscardin WJ, Adhere Scientific Advisory Committee SG, Investigators. Risk stratification for in-hospital mortality in acutely decompensated heart failure: classification and regression tree analysis. JAMA 2005;**293**(5):572-80.

22. Lee DS, Stitt A, Austin PC, Stukel TA, Schull MJ, Chong A, Newton GE, Lee JS, Tu JV. Prediction of heart failure mortality in emergent care: a cohort study. Ann Intern Med 2012;**156**(11):767-75.

23. O'Connor CM, Mentz RJ, Cotter G, Metra M, Cleland JG, Davison BA, Givertz MM, Mansoor GA, Ponikowski P, Teerlink JR, Voors AA, Fiuzat M, Wojdyla D, Chiswell K, Massie BM. The PROTECT in-hospital risk model: 7-day outcome in patients hospitalized with acute heart failure and renal dysfunction. Eur J Heart Fail 2012;**14**(6):605-12.

24. Cleland JG, Chiswell K, Teerlink JR, Stevens S, Fiuzat M, Givertz MM, Davison BA, Mansoor GA, Ponikowski P, Voors AA, Cotter G, Metra M, Massie BM, O'Connor CM. Predictors of postdischarge outcomes from information acquired shortly after admission for acute heart failure: a report from the Placebo-Controlled Randomized Study of the Selective A1 Adenosine Receptor Antagonist Rolofylline for Patients Hospitalized With Acute Decompensated Heart Failure and Volume Overload to Assess Treatment Effect on Congestion and Renal Function (PROTECT) Study. Circ Heart Fail 2014;**7**(1):76-87.

25. Lee DS, Austin PC, Rouleau JL, Liu PP, Naimark D, Tu JV. Predicting mortality among patients hospitalized for heart failure: derivation and validation of a clinical model. JAMA 2003;**290**(19):2581-7.

26. Felker GM, Leimberger JD, Califf RM, Cuffe MS, Massie BM, Adams KF, Jr., Gheorghiade M, O'Connor CM. Risk stratification after hospitalization for decompensated heart failure (OPTIME-CHF). J Card Fail 2004;**10**(6):460-6.

27. Khazanie P, Heizer GM, Hasselblad V, Armstrong PW, Califf RM, Ezekowitz J, Dickstein

K, Levy WC, McMurray JJ, Metra M, Tang WH, Teerlink JR, Voors AA, O'Connor CM, Hernandez AF, Starling R. Predictors of clinical outcomes in acute decompensated heart failure: Acute Study of Clinical Effectiveness of Nesiritide in Decompensated Heart Failure outcome models. Am Heart J 2015;**170**(2):290-7.

28. Sato N, Kajimoto K, Keida T, Mizuno M, Minami Y, Yumino D, Asai K, Murai K, Muanakata R, Aokage T, Sakata Y, Mizuno K, Takano T, Investigators T. Clinical features and outcome in hospitalized heart failure in Japan (from the ATTEND Registry). Circ J 2013;**77**(4):944-51.

29. Adams KF, Jr., Fonarow GC, Emerman CL, LeJemtel TH, Costanzo MR, Abraham WT, Berkowitz RL, Galvao M, Horton DP, Committee ASA, Investigators. Characteristics and outcomes of patients hospitalized for heart failure in the United States: rationale, design, and preliminary observations from the first 100,000 cases in the Acute Decompensated Heart Failure National Registry (ADHERE). Am Heart J 2005;**149**(2):209-16.

30. Harjola VP, Follath F, Nieminen MS, Brutsaert D, Dickstein K, Drexler H, Hochadel M, Komajda M, Lopez-Sendon JL, Ponikowski P, Tavazzi L. Characteristics, outcomes, and predictors of mortality at 3 months and 1 year in patients hospitalized for acute heart failure. Eur J Heart Fail 2010;**12**(3):239-48.

31. Park JJ, Choi DJ, Yoon CH, Oh IY, Jeon ES, Kim JJ, Cho MC, Chae SC, Ryu KH, Yoo BS, Kang SM, Oh BH, Kor HFR. Prognostic value of C-reactive protein as an inflammatory and N-terminal probrain natriuretic peptide as a neurohumoral marker in acute heart failure (from the Korean Heart Failure registry). Am J Cardiol 2014;**113**(3):511-7.

32. Shoaib A, Waleed M, Khan S, Raza A, Zuhair M, Kassianides X, Djahit A, Goode K, Wong K, Rigby A, Clark A, Cleland J. Breathlessness at rest is not the dominant presentation of patients admitted with heart failure. Eur J Heart Fail 2014;**16**(12):1283-91.

33. Takei M, Kohsaka S, Shiraishi Y, Goda A, Izumi Y, Yagawa M, Mizuno A, Sawano M, Inohara T, Kohno T, Fukuda K, Yoshikawa T, West Tokyo Heart Failure Registry I. Effect of estimated plasma volume reduction on renal function for acute heart failure differs between patients with preserved and reduced ejection fraction. Circ Heart Fail 2015;**8**(3):527-32.

34. Nagai T, Nishimura K, Honma T, Higashiyama A, Sugano Y, Nakai M, Honda S, Iwakami N, Okada A, Kawakami S, Kanaya T, Asaumi Y, Aiba T, Nishida Y, Kubota Y, Sugiyama D, Okamura T, Noguchi T, Kusano K, Ogawa H, Yasuda S, Anzai T, Na DEFi. Prognostic significance of endogenous erythropoietin in long-term outcome of patients with acute decompensated heart failure. Eur J Heart Fail 2016;**18**(7):803-13.

35. Geersing GJ, Bouwmeester W, Zuithoff P, Spijker R, Leeflang M, Moons KG. Search filters for finding prognostic and diagnostic prediction studies in Medline to enhance systematic reviews. PLoS One 2012;**7**(2):e32844.

36. Komajda M, Hanon O, Hochadel M, Lopez-Sendon JL, Follath F, Ponikowski P, Harjola

VP, Drexler H, Dickstein K, Tavazzi L, Nieminen M. Contemporary management of octogenarians hospitalized for heart failure in Europe: Euro Heart Failure Survey II. Eur Heart J 2009;**30**(4):478-86.

37. Sato N, Kajimoto K, Asai K, Mizuno M, Minami Y, Nagashima M, Murai K, Muanakata R, Yumino D, Meguro T, Kawana M, Nejima J, Satoh T, Mizuno K, Tanaka K, Kasanuki H, Takano T, Investigators A. Acute decompensated heart failure syndromes (ATTEND) registry. A prospective observational multicenter cohort study: rationale, design, and preliminary data. Am Heart J 2010;**159**(6):949-955 e1.

38. Hamaguchi S, Kinugawa S, Tsuchihashi-Makaya M, Goto D, Yamada S, Yokoshiki H, Tsutsui H. Characteristics, management, and outcomes for patients during hospitalization due to worsening heart failure-A report from the Japanese Cardiac Registry of Heart Failure in Cardiology (JCARE-CARD). J Cardiol 2013;**62**(2):95-101.

39. Makino J, Kohsaka S, Shiraishi Y, Inohara T, Goda A, Izumi Y, Yagawa M, Mizuno A, Fukuda K, Yoshikawa T. Application of the United States acute heart failure risk prediction model in Japanese patients; analysis from a contemporary multicenter registry. Int J Cardiol 2015;**195**:323-5.

40. Honda S, Nagai T, Sugano Y, Okada A, Asaumi Y, Aiba T, Noguchi T, Kusano K, Ogawa H, Yasuda S, Anzai T, Na DEFi. Prevalence, determinants, and prognostic significance of delirium in patients with acute heart failure. Int J Cardiol 2016;**222**:521-7.

41. Deubner N, Berliner D, Frey A, Guder G, Brenner S, Fenske W, Allolio B, Ertl G, Angermann CE, Stork S. Dysnatraemia in heart failure. Eur J Heart Fail 2012;**14**(10):1147-54.

42. Brisco MA, Coca SG, Chen J, Owens AT, McCauley BD, Kimmel SE, Testani JM. Blood urea nitrogen/creatinine ratio identifies a high-risk but potentially reversible form of renal dysfunction in patients with decompensated heart failure. Circ Heart Fail 2013;**6**(2):233-9.

43. Preiss DJ, Godber IM, Lamb EJ, Dalton RN, Gunn IR. The influence of a cooked-meat meal on estimated glomerular filtration rate. Ann Clin Biochem 2007;**44**(Pt 1):35-42.

44. Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013;**13**:33.

45. Ziaeian B, Kominski GF, Ong MK, Mays VM, Brook RH, Fonarow GC. National Differences in Trends for Heart Failure Hospitalizations by Sex and Race/Ethnicity. Circ Cardiovasc Qual Outcomes 2017;**10**(7).

46. Ouwerkerk W, Voors AA, Zwinderman AH. Factors influencing the predictive power of models for predicting mortality and/or heart failure hospitalization in patients with heart failure. JACC Heart Fail 2014;**2**(5):429-36.

Figure Legends

Figure 1-- Search strategy for the published HHF mortality prediction models. AUC = area under the curve; HHF = hospitalised due to worsening symptoms and signs of heart failure.

Figure 2-- Calibration plot. Predicted compared to observed mortality for the published models. The line shows the line of perfect calibration. Dots represent mortality for patients stratified by predicted mortality for each cohort. P-values are for Hosmer-Lemeshow goodness-of-fit tests. ADHERE = Acute Decompensated Heart Failure National Registry; ASCEND-HF = Acute Study of Clinical Effectiveness of Nesiritide in Decompensated Heart Failure; GWTG-HF = Get With the Guidelines-Heart Failure; OPTIME-CHF = Outcomes of a Prospective Trial of Intravenous Milrinone for Exacerbations of Chronic Heart Failure; OPTIMIZE-HF = Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure.

Figure 3-- Calibration plot for modified published prediction models substituting BUN with creatinine. Predicted compared to observed mortality for the published models. The line shows the line of perfect calibration. Dots represent mortality for patients stratified by predicted mortality for each cohort. P-values are for Hosmer-Lemeshow goodness-of-fit tests. Abbreviations as in Figure 2.

Supplemental Figure S 1-- Calibration plot, restricted only to patients without missing BUN (in the UK) and NYHA (in Japan) data. ADHERE = Acute Decompensated Heart Failure National Registry; ASCEND-HF = Acute Study of Clinical Effectiveness of Nesiritide in Decompensated Heart Failure; GWTG-HF = Get with the Guidelines-Heart Failure; OPTIME-CHF = Outcomes of a Prospective Trial of Intravenous Milrinone for Exacerbations of Chronic Heart Failure.

Table

Table 1 Published Mortality prediction models and the variables imputed on those models

Variables	**GWTG-HF**	**ADHERE**	**OPTIMIZE-HF**	**ASCEND-HF**	**OPTIME-CHF**
Reference	18	21	19	27	26
Published year	2010	2005	2008	2015	2004
Type of study	Registry	Registry	Registry	RCT	RCT
Number of patients	71,284	33,046	37,548	7,141	949
Number of Hospitals	287	263	259	398	80
Country	USA	USA	USA	#International	USA
Mortality	Hospital	Hospital	Hospital	30-day & 180-day	60-day
Age	✔	✔	✔	✔	✔
Systolic BP	✔	✔	✔	✔	✔
Sodium	✔		✔	✔	✔
***BUN**	✔	✔		✔	✔
SCr			✔		
Heart rate		✔			
***Black race**	✔				
COPD	✔				
LVSD			✔		
Primary cause for admission (HF/other)			✔		
NYHA Class IV				✔	✔

- **Modified GWTG-HF score**: *****Race was not available in the UK patients; because the assigned score was very low (0-3) and around 3% of population in England is British black, we assigned score "3" in all UK patients. BUN was replaced with creatinine
- **Modified ASCEND-HF score**: BUN replaced with creatinine
- **Modified OPTIME-CHF score**: BUN replaced with creatinine

****41.7% of patients in the UK had missing BUN value and hence was replaced with creatinine***

North America (44.9%), Europe (20.8%), Asia Pacific (24.9%), South America (9.4%)

ADHERE = Acute Decompensated Heart Failure National Registry; ASCEND-HF = Acute Study of Clinical Effectiveness of Nesiritide in Decompensated Heart Failure; BP = blood pressure; BUN = blood urea nitrogen; COPD = chronic obstructive pulmonary disease; GWTG-HF = Get With the Guidelines-Heart Failure; LVSD = left ventricular systolic dysfunction; NYHA = New York Heart Association; OPTIME-CHF = Outcomes of a Prospective Trial of Intravenous Milrinone for Exacerbations of Chronic Heart Failure; OPTIMIZE-HF = Organized Program to Initiate Lifesaving Treatment in Hospitalized Patients with Heart Failure; SCr = serum creatinine; UK = United Kingdom; USA = United States of America; RCT= Randomized clinical tria

Table 2 Baseline characteristics

Variable	% Missing UK/Japan	UK N=894	Japan N=3158	P-value
Age, years	0.1/0.0	76 (67-83)	78 (69-84)	<0.001
Male sex, n (%)	0/6.8	555 (62)	1744 (59)	0.13
BMI	18.1/8.6	28 (25-33)	23 (20-26)	<0.001
NYHA III or IV, n (%)	2.7/26.8	802 (92)	1889 (82)	<0.001
NYHA IV, n (%)		252 (29)	1077 (47)	<0.001
Systolic BP, mmHg	0.6/15.1	129 (110-149)	136 (118-159)	<0.001
Heart rate, /min	0.7/15.4	86 (72-106)	90 (72-110)	0.016
LVEF ≤45%, %	12.6/18.7	560 (72)	1377 (54)	<0.001
Comorbidities, n (%)				
Ischemic heart disease	0/0	433 (48)	778 (25)	<0.001
Diabetes mellitus	0.1/0.1	297 (33)	1187 (38)	0.017
Hypertension	0.1/0.2	512 (57)	2312 (73)	<0.001
COPD/Asthma	0/18.4	221 (25)	141 (5)	<0.001
Atrial fibrillation	0.6/18.4	374 (41)	1305 (51)	<0.001
Stroke	1.3/18.6	88 (10)	457 (18)	<0.001
Laboratory data				
Hemoglobin, g/dL	43.7/0.4	12.3 (10.9-13.6)	11.8 (10.1-13.3)	<0.001
Sodium, mEq/L	0.1/0.5	138 (135-141)	140 (137-142)	<0.001
Potassium, mEq/L	2.1/0.5	4.4 (4.0-4.8)	4.3 (3.9-4.7)	<0.001
BUN, mg/dL	41.7/0.5	24.7 (16.5-38.9)	22.7 (16.7-33.0)	0.004
Creatinine, mg/dL	0.5/0.6	1.3 (1.0-1.7)	1.1 (0.8-1.5)	<0.001
eGFR, mL/min/1.73m^2	0.5/7.3	52 (36-73)	66 (43-89)	<0.001
CRP, mg/dL	22.5/0.0	1.5 (0.5-3.5)	0.5 (0.1-1.8)	<0.001
Medications at discharge, n (%)	3158			
ACE-Is or ARBs	8.1/1.9	682 (83)	1945 (63)	<0.001
Beta blockers	8.1/1.9	640 (78)	2268 (73)	0.007
MRAs	7.2/7.7	424 (51)	962 (33)	<0.001
Diuretics	6.5/7.3	748 (89)	2070 (71)	<0.001

Digitalis	5.6/2.1	203 (24)	251 (8)	<0.001

Values are n (%) or median (interquartile range), unless otherwise indicated.

ACE-I = angiotensin converting enzyme; ARB = angiotensin receptor blocker; BMI = body mass index; BP = blood pressure; BUN = blood urea nitrogen; COPD = chronic obstructive pulmonary disease; CRP = C - reactive protein; eGFR = estimated glomerular filtration rate; LVEF = left ventricular ejection fraction; MRA = mineralocorticoid receptor antagonists; NYHA = New York Heart Association.

Table 3 Risk score and classification, predicted probability of death, duration of hospital stay, and actual mortality

Variable	UK N=894	Japan N=3158	P-value
Risk scores or log odds of mortality (for ADHERE)			
GWTG-HF for In-hospital death			
Score	45 ± 9	41 ± 8	<0.001
Predicted probability of death, %	8.1	6.6	-
Modified GWTG-HF for In-hospital death			
Modified score	48 ± 9	46 ± 9	<0.001
Predicted probability of death, %	4.9	4.1	-
ADHERE for In-hospital death			
Score	-3.1 ± 0.9	-3.4 ± 0.9	<0.001
Predicted probability of death, %	8.0	7.0	-
Modified ADHERE for In-hospital death			
Modified score	-3.9 ± 0.7	-4.0 ± 0.7	0.008
Predicted probability of death, %	4.8	4.6	-
OPTIMIZE-HF for In-hospital death			
Score	35 ± 8	34 ± 7	<0.001
Predicted probability of death, %	5.1	4.4	-
ASCEND-HF for 30-day death			
Score	3 (2-5)	3 (2-4)	<0.001
Predicted probability of death, %	5.6	5.2	-
Modified ASCEND-HF for 30-day death			
Modified score	3 (2-5)	3 (2-4)	<0.001
Predicted probability of death, %	5.6	5.0	-
OPTIME-CHF for 90-day death			
Score	186 ± 45	173 ± 44	<0.001
Predicted probability of death, %	16.1	13.8	-
Modified OPTIME-CHF for 90-day death			
Modified score	174 ± 38	164 ± 38	<0.001
Predicted probability of death, %	13.3	11.4	-
ASCEND-HF for 180-day death			

Predicted probability of death, %	23.7	20.7	-
Modified ASCEND-HF for 180-day death			
Predicted probability of death, %	20.5	17.9	-
Length of Hospital Stay, days	11 (7-18)	15 (10-24)	<0.001
Length of Hospital Stay >30 days, %	11.1	23.8	<0.001
Actual mortality			
In-hospital	43 (4.8)	80 (2.5)	<0.001
30-day	47 (5.7)	62 (2.6)	<0.001
90-day	118 (13.6)	138 (6.0)	<0.001
180-day	179 (20.7)	214 (9.5)	<0.001

Values are n (%), or mean ± SD, or median (interquartile range), unless otherwise indicated.

Abbreviations as in Table 1.

Table 4 Discriminative value of risk models

Models	Derivation cohort C-statistic (95% CI)	UK C-statistic (95% CI)	Japan C-statistic (95% CI)
In-hospital			
GWTG-HF	0.75	0.669 (0.594 - 0.737)	0.771 (0.710 - 0.832)
Modified GWTG-HF		0.722 (0.653 - 0.791)	0.758 (0.696 - 0.820)
ADHERE	0.759	0.663 (0.590 - 0.735)	0.755 (0.694 - 0.816)
Modified ADHERE		0.695 (0.622 - 0.768)	0.689 (0.624 - 0.755)
OPTIMIZE-HF	0.753 (0.741-0.765)	0.752 (0.683 - 0.810)	0.767 (0.702 - 0.832)
30-day			
ASCEND-HF	0.75	0.619 (0.521 - 0.717)	0.750 (0.677 - 0.822)
Modified ASCEND-HF		0.661 (0.592 - 0.730)	0.731 (0.659 - 0.802)
90-day			
OPTIME-CHF	0.76	0.673 (0.607 - 0.739)	0.751 (0.706 - 0.796)
Modified OPTIME-CHF		0.701 (0.650 - 0.753)	0.730 (0.683 - 0.777)
180-day			
ASCEND-HF	0.70	0.689 (0.637 - 0.742)	0.748 (0.712 - 0.785)
Modified ASCEND-HF		0.689 (0.647 - 0.731)	0.734 (0.697 - 0.770)

Other abbreviations as in Table 1.

Figure

Figure 1

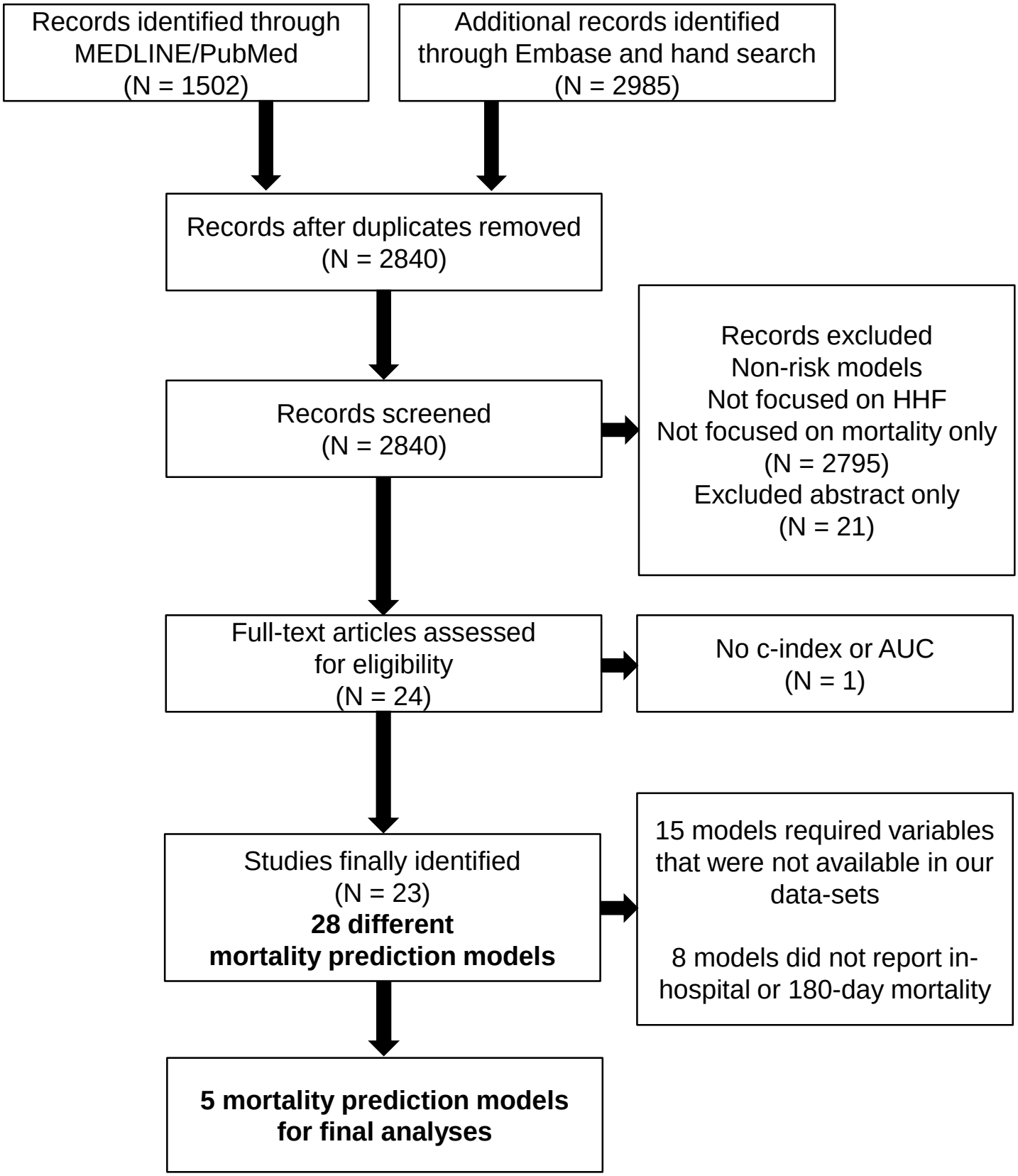


Figure 2

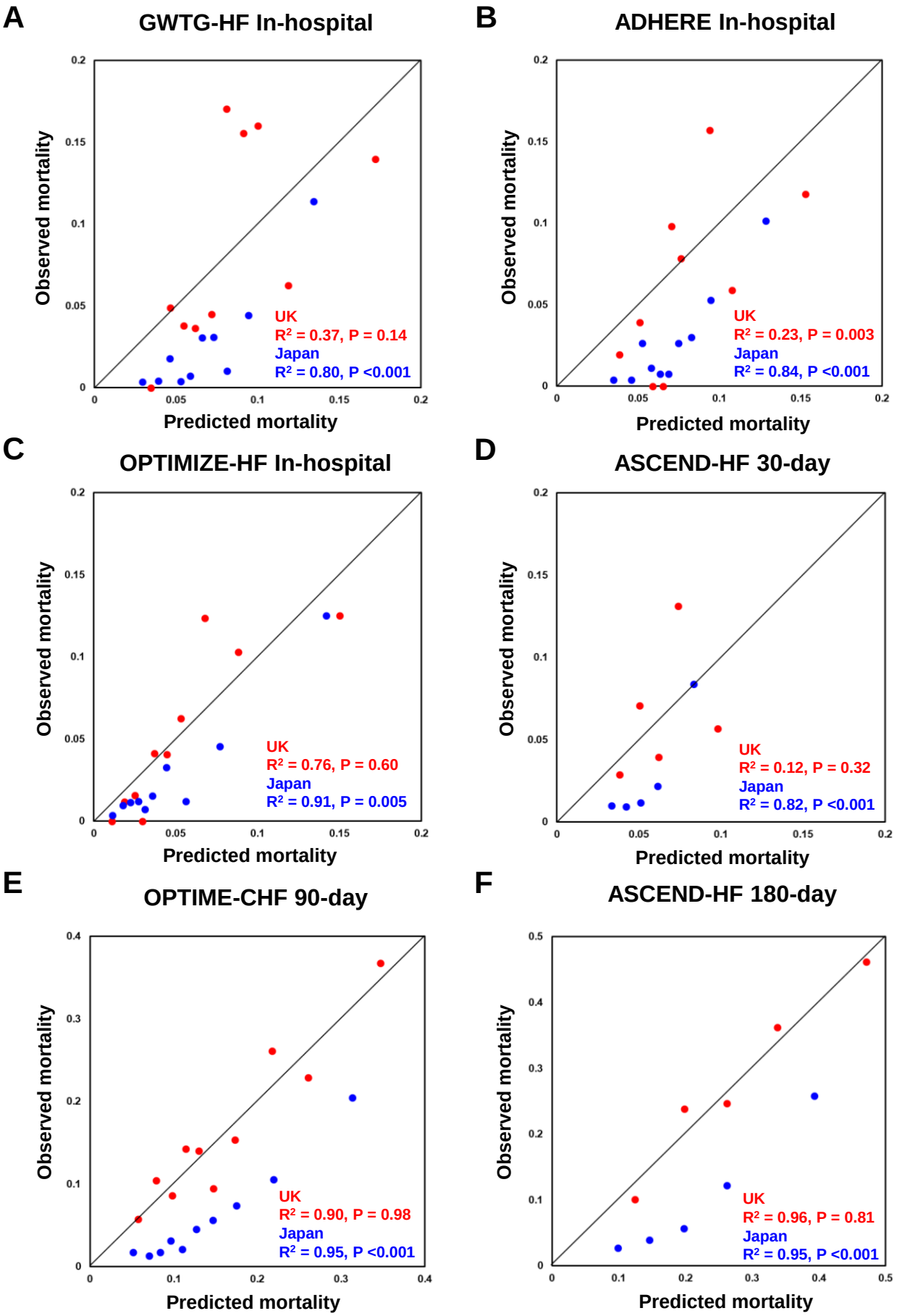
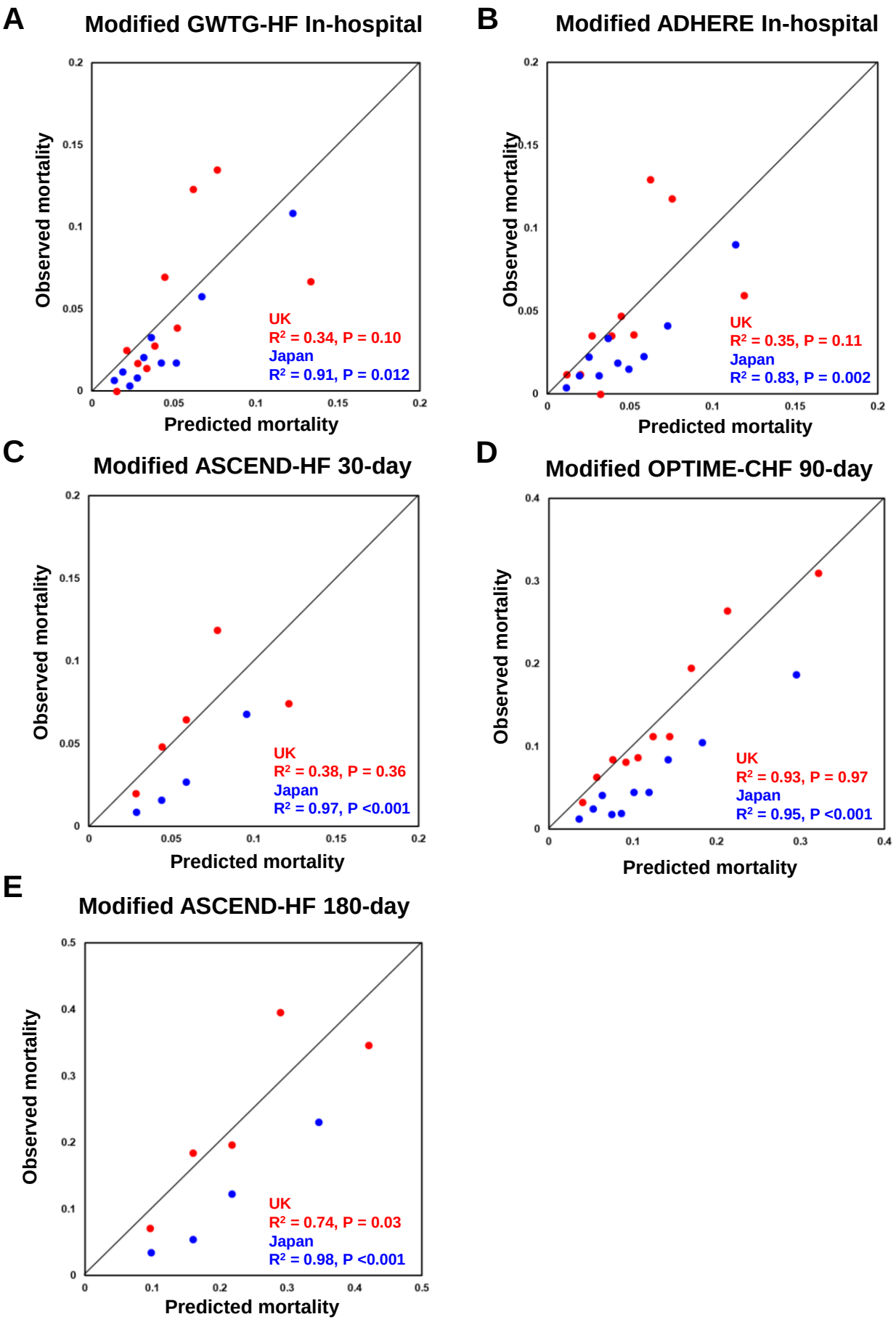

A
GWTG-HF In-hospital
Observed mortality
Predicted mortality
UK
R² = 0.37, P = 0.14
Japan
R² = 0.80, P <0.001
B
ADHERE In-hospital
Observed mortality
Predicted mortality
UK
R² = 0.23, P = 0.003
Japan
R² = 0.84, P <0.001
C
OPTIMIZE-HF In-hospital
Observed mortality
Predicted mortality
UK
R² = 0.76, P = 0.60
Japan
R² = 0.91, P = 0.005
D
ASCEND-HF 30-day
Observed mortality
Predicted mortality
UK
R² = 0.12, P = 0.32
Japan
R² = 0.82, P <0.001
E
OPTIME-CHF 90-day
Observed mortality
Predicted mortality
UK
R² = 0.90, P = 0.98
Japan
R² = 0.95, P <0.001
F
ASCEND-HF 180-day
Observed mortality
Predicted mortality
UK
R² = 0.96, P = 0.81
Japan
R² = 0.95, P <0.001

Figure 3


A
Modified GWTG-HF In-hospital
Observed mortality
Predicted mortality
UK
R² = 0.34, P = 0.10
Japan
R² = 0.91, P = 0.012
B
Modified ADHERE In-hospital
Observed mortality
Predicted mortality
UK
R² = 0.35, P = 0.11
Japan
R² = 0.83, P = 0.002
C
Modified ASCEND-HF 30-day
Observed mortality
Predicted mortality
UK
R² = 0.38, P = 0.36
Japan
R² = 0.97, P <0.001
D
Modified OPTIME-CHF 90-day
Observed mortality
Predicted mortality
UK
R² = 0.93, P = 0.97
Japan
R² = 0.95, P <0.001
E
Modified ASCEND-HF 180-day
Observed mortality
Predicted mortality
UK
R² = 0.74, P = 0.03
Japan
R² = 0.98, P <0.001

Supplementary Figures