A registry-based algorithm to predict ejection fraction in patients with heart failure

Alicia Uijl1,2,3*, Lars H. Lund1,4, Ilonca Vaartjes2, Jasper J. Brugts5, Gerard C. Linssen6, Folkert W. Asselbergs3,7,8, Arno W. Hoes2, Ulf Dahlström9, Stefan Koudstaal3,7 and Gianluigi Savarese1

1Division of Cardiology, Department of Medicine, Karolinska Institutet, Stockholm, Sweden; 2Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; 3Health Data Research UK London, Institute for Health Informatics, University College London, London, UK; 4Heart and Vascular Theme, Karolinska University Hospital, Stockholm, Sweden; 5Department of Cardiology, Erasmus University Medical Center, Thoraxcenter, Rotterdam, The Netherlands; 6Department of Cardiology, Hospital Group Twente, Almelo and Hengelo, The Netherlands; 7Department of Cardiology, Division Heart and Lungs, University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands; 8Institute of Cardiovascular Science, Faculty of Population Health Sciences, University College London, London, UK; 9Department of Cardiology and Department of Health, Medicine and Caring Sciences, Linkoping University, Linköping, Sweden

Abstract

Aims Left ventricular ejection fraction (EF) is required to categorize heart failure (HF) [i.e. HF with preserved (HFpEF), mid-range (HFmEF), and reduced (HFrEF)] but is often not captured in population-based cohorts or non-HF registries. The aim was to create an algorithm that identifies EF subphenotypes for research purposes.

Methods and results We included 42,061 HF patients from the Swedish Heart Failure Registry. As primary analysis, we performed two logistic regression models including 22 variables to predict (i) EF ≥ 50% and (ii) EF > 40%. In the secondary analysis, we performed a multivariable multinomial analysis with 22 variables to create a model for all three separate EF subphenotypes: HFrEF vs. HfMfEF vs. HFpEF. The models were validated in the database from the CHECK-HF study, a cross-sectional survey of 10,627 patients from the Netherlands. The C-statistic (discrimination) was 0.78 [95% confidence interval (CI) 0.77–0.78] for EF ≥ 50% and 0.76 [95% CI 0.75–0.76] for EF ≥ 40%. Similar results were achieved for HFrEF and HFpEF in the multinomial model, but the C-statistic for HFmEF was lower: 0.63 [95% CI 0.63–0.64]. The external validation showed similar discriminative ability to the development cohort.

Conclusions Routine clinical characteristics could potentially be used to identify different EF subphenotypes in databases where EF is not readily available. Accuracy was good for the prediction of HfEF and HFpEF but lower for HFmEF. The proposed algorithm enables more effective research on HF in the big data setting.

Keywords Electronic health records; Heart failure; Ejection fraction; Prediction; HFrEF; HFmEF; HFpEF

Introduction

Left ventricular ejection fraction (EF) is used in heart failure (HF) for diagnosis, characterization and treatment selection and is a key inclusion criterion for HF trials.1 Current European guidelines classify HF according to EF as HF with preserved EF (HFpEF; EF ≥ 50%), HF with mid-range EF (HFmEF; EF = 40–49%), and HF with reduced EF (HFrEF; EF < 40%).2

Electronic health records (EHRs) provide an abundance of routine clinical care data, which may contribute to assess quality of care and uncover the current unmet needs in HF, i.e. identifying underuse of evidence-based therapies and reasons for undertreatment in order to implement care.3–5 Furthermore, phenotyping real-world HF patients could facilitate the development of new treatments or the establishment of new uses of existing treatments and may also help in designing of and pre-screening for randomized trials in all EF categories. However, EHRs frequently lack readily available phenotypic information that is needed to discern relevant subphenotypes.6–9 In the case of HF, EF is often missing or not documented in EHRs, thereby preventing analyses focusing on specific EF subphenotypes and limiting EHRs use in HF research.
Previous algorithms have been developed for the purpose of identifying EF subphenotypes (i.e. HfpEF vs. HFmrEF vs. HFrEF) in routine care data using International Classification of Diseases (ICD) codes, but none have considered routine clinical information that may be relevant for EF prediction in trials data sets, registries, and EHRs.

Therefore, we aimed to develop and validate algorithms to discern HFrEF, HFmrEF, and HfpEF subphenotypes using two representative, large, contemporary HF registries.

Methods

Development cohort

The Swedish Heart Failure Registry (SwedeHF) has been previously described. Briefly, it was created in 2000 and broadly implemented throughout Sweden by 2003. The only inclusion criterion is clinician-judged HF. Patients are registered at discharge from hospital or after outpatient clinic visit on a web-based care report form and entered into the database (managed by Uppsala Clinical Research Center, Uppsala, Sweden).

All permanent residents in Sweden have unique personal identification numbers that allows linking of disease-specific health registries, governmental health, and statistical registries. For the current analysis, we linked SwedeHF to the National Patient Registry, which provided more data on baseline comorbidities.

In this study, we included 42 061 patients with known EF registered between 11 May 2000 and 31 December 2012. In SwedeHF, EF is recorded as a categorical variable, i.e. <30%, 30–39%, 40–49%, and ≥50%. We defined HFrEF as EF <40%, HFmrEF as EF between 40% and 49%, and HfpEF as EF ≥50%. The study flow diagram is reported in Supporting Information, Figure S1A.

Validation cohort

The CHECK-HF (Chronic Heart Failure ESC-guideline based Cardiology Practice Quality project) registry is a cross-sectional collection of 10 910 unselected patients with the diagnosis of chronic HF treated at outpatient HF clinics (96%) of 34 Dutch hospitals or encountered at the general cardiology outpatient clinic of the same hospitals (4%) between September 2013 and September 2016. Inclusion criteria for this study were 18 years of age or older and known EF (n = 10 627). EF was recorded as a continuous variable but recoded to HFrEF <40%, HFmrEF = 40–49%, and HfpEF ≥50%. The study flow diagram is reported in Supporting Information, Figure S1B.

Statistical methods

Baseline characteristics and missing data

Patient characteristics were summarized by HF subphenotype as mean (SD) or median (interquartile range [IQR]) for continuous variables and percentages for categorical variables. Multiple imputation using the mice algorithm in the statistical software package R was used to impute missing data for the variables included in the models. Supporting Information, Table S2 shows the variables included in the multiple imputation models and the amount of missing records in the SwedeHF data set. We generated 10 imputed data sets, and analyses were performed on each imputed data set separately. The results were then pooled using Rubin’s rules. All the analyses, except for descriptive statistics, were performed on imputed data.

Development of predictive models

In the primary analysis, we used multivariable logistic regression to fit two different predictive models: one for ≥50% (HfpEF) vs. EF <50% (HFrEF and HFmrEF) and one for EF <40% (HFrEF) vs. ≥40% (HFmrEF and HfpEF). For the secondary analysis, we used a multinomial logistic model to separately predict HfpEF, HFmrEF, and HFrEF (HFrEF was used as reference).

We screened several sources of EHR for commonly available variables to assess as potential predictors of EF subphenotypes in our analyses, and we selected the following: age, sex, clinical characteristics [NT pro B-type natriuretic peptide (NT-proBNP), New York Heart Failure Association (NYHA) class, mean arterial pressure, heart rate, body mass index (BMI), and estimated glomerular filtration rate (eGFR)], comorbidities [history of ischaemic heart disease, atrial fibrillation, chronic obstructive pulmonary disease (COPD), diabetes, hypertension, anaemia, cancer in the previous 3 years, and valvular disease], and treatments [device therapy (implantable cardioverter defibrillator or cardiac resynchronization therapy), renin–angiotensin system (RAS) inhibitors, beta-blockers, diuretics, mineralocorticoid receptor antagonist (MRA), and digoxin].

Variance inflation factor was used to test for multicollinearity among predictors. If a pair of predictors was highly correlated (variance inflation factor > 10), we included only one of the predictors in the multivariable model. We performed backward selection on the multivariable model based on Akaike’s information criterion to regress the full model towards the final model. Predicted probability threshold cut-offs for the prediction of EF subphenotypes were investigated to maximize accuracy, sensitivity, and specificity of the model.

Model discrimination

Area under the receiver operating curves were used to discern model discrimination. The C-statistic was used to assess model performance. For the secondary analysis, i.e. multinomial models, discrimination and calibration were calculated with a one-vs.-rest approach. The outcome for each EF category j was dichotomized, i.e. HFrEF vs. HFmrEF and HfpEF.
The C-statistic was then obtained by evaluating the predicted risk of EF category $j$ vs. the predicted risk of the remaining categories.\textsuperscript{15,16} Observed vs. predicted plots were created to visually assess model calibration. We externally validated the models in the CHECK-HF registry.

**Sensitivity analysis**

In a sensitivity analysis, we simplified the models by excluding the clinical variables (NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI, and eGFR) and therefore only investigated demographics, comorbidities, and treatments. This was done because many EHRs, such as claim databases, include categorical data but not clinical variables that are often continuous (e.g. chronic kidney disease rather than eGFR) or ordinal (e.g. NYHA class).

In further sensitivity analyses, we excluded only NT-proBNP and then NT-proBNP + NYHA class because both are HF specific variables that are not always recorded in EHRs. All statistical analyses were performed in R software version 3.5.1.

**Results**

**Baseline characteristics**

Baseline patient characteristics are summarized in Table 1. In the SwedeHF cohort, 56% of patients had HFrEF, 21% HFmrEF, and 23% HFpEF. Overall, HFpEF patients were older,

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Baseline characteristics of the SwedeHF cohort</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HFrEF</td>
</tr>
<tr>
<td>N</td>
<td>23402 (55.6%)</td>
</tr>
<tr>
<td>Demographics</td>
<td></td>
</tr>
<tr>
<td>Age [years, mean (SD)]</td>
<td>71.66 (12.33)</td>
</tr>
<tr>
<td>Sex [female (%)]</td>
<td>6745 (28.8)</td>
</tr>
<tr>
<td>Heart failure measurements</td>
<td></td>
</tr>
<tr>
<td>NYHA class [Class III/IV (%)]</td>
<td>8187 (45.8)</td>
</tr>
<tr>
<td>NT-proBNP [= &gt;median (%)]</td>
<td>4015 (55.8)</td>
</tr>
<tr>
<td>Clinical variables</td>
<td></td>
</tr>
<tr>
<td>Systolic blood pressure [mean (SD)]</td>
<td>124.39 (20.49)</td>
</tr>
<tr>
<td>Diastolic blood pressure [mean (SD)]</td>
<td>73.38 (12.26)</td>
</tr>
<tr>
<td>MAP [≥90 mmHg (%)]</td>
<td>11964 (51.8)</td>
</tr>
<tr>
<td>Heart rate [≥70 BPM (%)]</td>
<td>13244 (60.5)</td>
</tr>
<tr>
<td>BMI (%)</td>
<td></td>
</tr>
<tr>
<td>&lt;18.5 kg/m$^2$</td>
<td>336 (3.1)</td>
</tr>
<tr>
<td>18.5–24.9 kg/m$^2$</td>
<td>4369 (40.1)</td>
</tr>
<tr>
<td>25–29.9 kg/m$^2$</td>
<td>3896 (35.8)</td>
</tr>
<tr>
<td>≥30 kg/m$^2$</td>
<td>2290 (21.0)</td>
</tr>
<tr>
<td>eGFR (%)</td>
<td></td>
</tr>
<tr>
<td>≥90 mL/min/1.73 m$^2$</td>
<td>2761 (11.8)</td>
</tr>
<tr>
<td>60–89.9 mL/min/1.73 m$^2$</td>
<td>9630 (41.3)</td>
</tr>
<tr>
<td>30–59.9 mL/min/1.73 m$^2$</td>
<td>9273 (39.7)</td>
</tr>
<tr>
<td>&lt;30 mL/min/1.73 m$^2$</td>
<td>1669 (7.2)</td>
</tr>
<tr>
<td>Anaemia [Yes (%)]</td>
<td>7348 (31.4)</td>
</tr>
<tr>
<td>Revascularised [Yes (%)]</td>
<td>7536 (32.2)</td>
</tr>
<tr>
<td>Comorbidities</td>
<td></td>
</tr>
<tr>
<td>Atrial fibrillation [Yes (%)]</td>
<td>11936 (51.0)</td>
</tr>
<tr>
<td>COPD [Yes (%)]</td>
<td>3710 (15.9)</td>
</tr>
<tr>
<td>Diabetes [Yes (%)]</td>
<td>6257 (26.7)</td>
</tr>
<tr>
<td>Hypertension [Yes (%)]</td>
<td>12670 (54.1)</td>
</tr>
<tr>
<td>Ischaemic heart disease [Yes (%)]</td>
<td>12994 (57.8)</td>
</tr>
<tr>
<td>Myocardial infarction [Yes (%)]</td>
<td>9975 (42.6)</td>
</tr>
<tr>
<td>Peripheral artery disease [Yes (%)]</td>
<td>2277 (9.7)</td>
</tr>
<tr>
<td>Cancer previous 3 years [Yes (%)]</td>
<td>2896 (12.4)</td>
</tr>
<tr>
<td>Valvular disease [Yes (%)]</td>
<td>5335 (23.4)</td>
</tr>
<tr>
<td>Therapy</td>
<td></td>
</tr>
<tr>
<td>RAS inhibitor [Yes (%)]</td>
<td>21037 (90.4)</td>
</tr>
<tr>
<td>Beta-blocker [Yes (%)]</td>
<td>21045 (90.3)</td>
</tr>
<tr>
<td>Loop diuretic [Yes (%)]</td>
<td>18534 (79.6)</td>
</tr>
<tr>
<td>MRA [Yes (%)]</td>
<td>7591 (32.7)</td>
</tr>
<tr>
<td>Digoxin [Yes (%)]</td>
<td>4092 (17.6)</td>
</tr>
<tr>
<td>Device therapy [Yes (%)]</td>
<td>1421 (6.1)</td>
</tr>
</tbody>
</table>

BMI, body mass index; BPM, beats per minute; COPD, chronic obstructive pulmonary disease; eGFR, estimated Glomerular filtration rate; HFmrEF, heart failure with mid-range ejection fraction; HFpEF, heart failure with preserved ejection fraction; HFrEF, heart failure with reduced ejection fraction; MAP, mean arterial pressure; mean (SD), mean (standard deviation); MRA, mineralocorticoid receptor antagonist; NT-proBNP, N-terminal pro b-type natriuretic peptide; NYHA class, New York Heart Association class; RAS inhibitor, renin angiotensin system inhibitor.

ESC Heart Failure ; : DOI: 10.1002/ehf2.12779
more likely female patients, and had higher blood pressure and BMI. Generally, comorbidities were more likely observed in HfPEF compared with HfREF and HFmrEF, except for history of myocardial infarction, which was considerably more common in HfREF and HFmrEF. HfREF but also HFmrEF patients were more likely to receive RAS inhibitors, beta-blockers, MRAs, and device therapy compared with HfPEF patients, while diuretics were more often prescribed in HfPEF patients. Baseline characteristics of the external validation cohort are summarized in Supporting Information, Table S2. Similar characteristics were overall observed in the CHECK-HF population and its subphenotypes. However, in the CHECK-HF vs. SwedeHF cohort, there were slightly less HFmrEF (15%) and HfPEF (21%) patients but more HfREF (64%) patients. HfREF patients were slightly more likely female patients (35% vs. 29%, respectively) and had lower NYHA class (28% vs. 46% NYHA class III/IV, respectively). Regardless of the EF subphenotype, in the CHECK-HF vs. SwedeHF cohort, there were less patients with anaemia and cardiovascular comorbidities such as hypertension, ischaemic heart disease, atrial fibrillation, and valvular disease. CHECK-HF patients were more likely to receive MRAs and device therapy but less likely to receive RAS inhibitors and beta-blockers compared with those in SwedeHF, regardless of the EF subphenotype.

Prediction models

Primary analysis

The model predicting EF $\geq 50\%$ vs. $<50\%$ is presented in Figure 1. The strongest predictors [those with an odds ration (OR) $> 1.5$] for EF $\geq 50\%$ were older age, female sex, hypertension, anaemia, and atrial fibrillation. Device therapy, use of RAS inhibitors, and higher NT-proBNP levels had the strongest association with EF $< 50\%$ (OR $< 0.5$). The model discriminated well, with a C-statistic of 0.775 [95% confidence interval (95% CI) 0.770–0.780] (Figure 3A). There was a slight overestimation for the predicted probabilities between 0.4 and 0.6 (Figure 4A). With a predicted probability threshold of 0.21, we maximized the sensitivity and specificity of predicting EF $\geq 50\%$, while a higher threshold of 0.44 led to a higher overall accuracy and higher specificity to predict EF $< 50\%$ (Supporting Information, Table S3).

Comparable results were observed for the model predicting EF $\geq 40\%$ vs. $<40\%$, with older age and female sex as strongest predictors for EF $\geq 40\%$ (Figure 2). Furthermore, BMI $\geq 30$ kg/m$^2$, atrial fibrillation, hypertension, and anaemia were strong predictors for EF $\geq 40\%$ (OR $> 1.5$), while device therapy, RAS inhibitors, and higher NT-proBNP levels were the strongest predictors for EF $<40\%$ (OR $< 0.5$). The discrimination of this model was good, with a C-statistic of 0.757.
(95% CI 0.752–0.763) (Figure 3B) and slight underestimation and overestimation in the lower and higher ends of the predicted probabilities (Figure 4B). Predicted probability thresholds to maximize overall accuracy or sensitivity + specificity was similar, with cut-offs of 0.48 and 0.45, respectively (Supporting Information, Table S3).

Secondary analysis
The results from the multinomial model are shown in Figure 2. HFrefEF was the reference category. Compared with HFrefEF, older age, female sex, higher BMI and atrial fibrillation were the strongest predictors for HFrEF. Predictors for HFpEF were similar to those for HFrEF, but the associations were much stronger. C-statistics according to the one-vs.-rest approach for HFrEF and HFpEF were similar to the logistic models for EF ≥40% or EF ≥50% in the primary analysis, 0.758 (95% 0.754–0.763) and 0.775 (95% 0.770–0.780), respectively (Figure 3C). However, the discriminative performance for predicting HFrEF was only moderate, with a C-statistic of 0.633 (95% CI 0.627–0.640). Model calibration was not optimal (Figure 4C). Overall accuracy was much lower for the multinomial model than for the primary analyses, with an accuracy of 58.1–60.8% (Supporting Information, Table S3).

External validation
Models were externally validated in the CHECK-HF data set, with good discriminative performance that was comparable with the development cohort, and the EF ≥50% models performing best with a C-statistic of 0.728 (0.724–0.731) for the main model (Supporting Information, Table S4).

Sensitivity analyses
We performed sensitivity analyses to investigate simpler models, i.e. excluding clinical characteristics (NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI, and eGFR) (Supporting Information, Tables S5, S6, and S12) as well as models excluding only NT-proBNP (Supporting Information, Tables S7, S8, and S12) and models excluding NT-proBNP and NYHA (Supporting Information, Table S9, S10, and S13). The models had lower but good discriminative ability for EF ≥50% vs. <50% (Supporting Information, Figures S2, S4, and S6), with a C-statistic for the simple model of 0.737 (95% CI 0.732–0.743), 0.753 (95% CI 0.748–0.759) for the model without NT-proBNP, and 0.750 (95% CI 0.744–0.755) for the model without NT-proBNP and NYHA. This was similar for the logistic model predicting EF ≥40% vs. <40%, with a C-statistic of 0.703 (95% CI 0.698–0.708) for the simpler model, 0.734 (95% CI 0.729–0.739) for the logistic model excluding NT-proBNP, and 0.721 (95% CI 0.716–0.726) for the model excluding NT-proBNP and NYHA (Supporting Information, Figures S3, S5, and S7). Likewise, HFrEF and HFpEF at the multinomial analysis had good discriminative ability, while predicting HFrEF was only moderate (Supporting Information, Figures S8–S10).

Figure 2. Multinomial prediction model predicting HFrEF or HFpEF with HFrEF as reference category. BMI, body mass index; BPM, beats per minute; COPD, chronic obstructive pulmonary disease; eGFR, estimated Glomerular filtration rate; HFrEF, heart failure with mid-range ejection fraction; HFpEF, heart failure with preserved ejection fraction; MAP, mean arterial pressure; MRA, mineralocorticoid receptor antagonist; NT-proBNP, N-terminal pro b-type natriuretic peptide; NYHA class, New York Heart Association class; RAS inhibitor, renin angiotensin system inhibitor.
We externally validated these sensitivity analyses in the CHECK-HF data set, with similar discriminative performances as in the development cohort (Supporting Information, Table S4).

**Discussion**

EHRs and routine clinical care data represent a great potential resource for HF research. While these databases provide for large samples sizes ensuring generalizability and many clinically relevant variables, the main limitation is often the depth of phenotypic information required to identify and investigate specific HF subphenotypes. Currently, EF is the key to phenotype HF patients and is used for treatment selection in clinical practice and as inclusion criterion in HF trials. Moreover, as shown in numerous previous studies, patients with different EF subphenotypes have different risk profiles, disease trajectories, and outcomes. Absence of readily available EF measurements limits research on HF in routine EHR data. Several natural language processing models could be used to extract data on left ventricular systolic function reported as free text in EHR. For those instances that this information is not available, simple prediction models for EF might be used to gain more knowledge on HF phenotypic information in EHRs, claim databases, trials, and large cohorts. With recent data on angiotensin-receptor-Neprilysin inhibitors and potentially emerging data on sodium/glucose cotransporter 2 inhibitors in HF, the use of these drugs may be expanded. It would be important for regulators, payers, and health systems to be able to use EF prediction models to assess implications of these new drugs in their own health care systems and databases.

We hereby propose prediction models that could be used to infer EF category in secondary care HF patients based on patients’ characteristics for research purposes. Our models discriminated well, especially for HFrEF and HFpEF, while predicting HFmrEF was more challenging.

Two previous studies aimed to create algorithms to predict EF category in HF patients. Bovitz et al. realized a predictive model for EF based on ICD-9 codes for systolic and...
diastolic HF in 2714 patients encountered in a single centre. The area under the curve for this model was 0.821 and had a predicted probability threshold cut-off for EF of 43.5%.10 The main limitation was generalizability. Indeed, no external validation was performed, and this study enrolled a small cohort of patients from only one centre, whereas ICD coding practice is highly varying from one centre to another. Furthermore, this model did not incorporate clinical or laboratory data such as blood pressure, eGFR, or NT-proBNP. A predictive model from Desai et al. included 11 073 patients (of which 7105 patients are in the development cohort) and aimed to predict HFrEF, HFmrEF, or HfPEF as well as with EF< or ≥45% in patients with known EF from a centre referring to Medicare (claim database).11 The discriminative performance varied between 0.84 and 0.88. This model was externally validated in a cohort of patients from a different hospital but still limited to Medicare patients only.

Compared with previous models that have been developed to be mainly applied to claim data, our model, which considers also clinically relevant variables, can be used as well in clinical cohorts or trials where HF is diagnosed at baseline but EF is not collected.25 Furthermore, we have developed predicted probability thresholds to optimize accuracy or sensitivity and specificity that can guide researchers in classifying patients based on our models.

We created prediction models for HFrEF, HFmrEF, and HfPEF as well as for EF ≥40% vs <40% and EF ≥50% vs. <50% in SwedeHF. Our models had good performance, with the lowest C-statistic 0.633 for HFmrEF in the multinomial model and the highest performance for the EF ≥50% model with a C-statistic of 0.775. The lower C-statistic for HFmrEF may be explained by the heterogeneity that characterizes this subphenotype,16,26,27 with a large proportion of patients having transitioning EF for different reasons (e.g. atrial fibrillation and ischaemic heart disease) that may make EF prediction more challenging.28 Most trials use EF 40% or 50% as cut-offs for enrolment, and we provided models to identify patients based on these cut-offs (i.e. EF ≥40% vs <40% and EF ≥50% vs. <50%). If a trial or other research programme wishes to specifically select HFrEF, HfPEF, or HFmrEF patients, our models to identify the specific subphenotype could be applied, albeit that the area under the curve was worse (0.633) than for the dichotomous models (0.775 and 0.757, respectively).

Similar to the binary model by Desai et al.,11 male sex, implantable devices, and use of ACE inhibitors, beta-blockers, and MRAs predicted HFrEF in both models using an EF of 40% and 50% as cut-offs, while anaemia, valvular disease, obesity, and hypertension were predictive of HfPEF. Out of the comorbidities we included in our model, only ischaemic heart disease was predictive for HFrEF or EF <50%. This is comparable with what is known from recent studies, i.e. HfPEF is more related to ageing, female sex, and comorbidities, while HFrEF (and HFmrEF) are more likely to be associated with ischaemic heart disease.17–20 The main variables associated with HFrEF were medication use and variables associated with worsening or symptomatic HF, such as higher NYHA class and higher NT-proBNP levels. While medication use is not directly involved in the pathophysiology of any HF subphenotype, it is still helpful as a marker reflecting clinician decisions that in turn reflect EF. Interestingly, only severe renal disease (eGFR < 30 mL/min/1.73 m<sup>2</sup>) was associated with HFrEF, while mildly reduced kidney function was not associated with either EF subphenotype.

Strengths and limitations

SwedeHF and CHECK-HF are both large, unselected, contemporary HF cohorts, collecting data on demographics, clinical characteristics, biomarkers, medication use, and, notably, EF measurements. A strength of our analysis is that we were able to externally validate our models from SwedeHF in an independent sample with good discriminative performance (CHECK-HF). Furthermore, SwedeHF data was collected between 2000-2012, while the CHECK-HF registry was conducted between 2013-2016, indicating that the model performs well over time. However, there are also several limitations which need to be mentioned. First, EF is collected as a categorical variable in SwedeHF; therefore, we were unable to investigate linear associations between predictors and EF. However, clinical guidelines and trials use EF categories as well and would not be improved by linear information. Based on our models, it remains difficult to classify HFmrEF, which may be misclassified as HFrEF or HfPEF, and, therefore, we rather suggest using the models pooling HFmrEF with HfPEF or HFrEF. Second, many of the HF therapies were predictive for HFrEF/HfPEF and thus, when applying our models, we suggest considering the use of medications for 3–6 months after the initial HF diagnosis to allow for optimizing therapies and reflection of clinician decision making. Third, the inclusion criterion for SwedeHF is clinician-judged HF, which differs from the ICD definition of HF in EHRs and thus our model should be further evaluated and validated in an EHR setting. Finally, repeated measurements of clinical characteristics (e.g. NYHA class, blood pressure, etc.) and EF are limited in SwedeHF and thus we could not assess how sensitive our model is to reclassify the patient EF subphenotype based on changes in clinical measurements.

Conclusions

We created an algorithm based on patient demographics, clinical characteristics and use of treatments to identify EF subphenotypes in HF patients without an available EF assessment. Accuracy was good for the prediction of HfPEF and
HFrEF but lower for HfmrEF, perhaps due to the heterogeneity that characterizes this subphenotype. Our model could significantly support more effective research in the ‘big data’ setting.

**Conflict of interest**

A. U., I. V., J. B., G. L., F. A., A. H., and S. K. have nothing to disclose. U. D. reports grants from AstraZeneca, Boehringer Ingelheim, and personal fees from Novartis Pharma, AstraZeneca, and Amgen, outside the submitted work. G. S. reports grants and personal fees from Vifor, grants and non-financial support from Boehringer Ingelheim, personal fees from SPA, grants from MSD, grants and personal fees from AstraZeneca, and personal fees from Roche, Medtronic, and Cytokinetics outside the submitted work. L. H. L. reports personal fees from Merck, grants and personal fees from Boehringer Ingelheim, personal fees from Sanofi, grants and personal fees from Vifor-Fresenius, personal fees from AstraZeneca, grants and personal fees from Relypsa, personal fees from Bayer, grants from Boston Scientific, grants and personal fees from Novartis, personal fees from Pharmacosmos, personal fees from Abbott, grants and personal fees from Mundipharma, and personal fees from Medscape, outside the submitted work.

**Funding**

The Swedish Heart Failure Registry is funded by the Swedish National Board of Health and Welfare, the Swedish Association of Local Authorities and Regions, the Swedish Society of Cardiology, and the Swedish Heart-Lung Foundation. Servier, the Netherlands, partially funded the inclusion of data and software program for CHECK-HF. The CHECK-HF steering committee (J. B., G. L., H. P. B., and A. H.) received no funding for this project. The current study was initiated by the authors and was designed, conducted, interpreted, and reported independently of the sponsor. This work has received support from the EU/EFPIA Innovative Medicines Initiative 2 Joint Undertaking BigData@Heart (grant no. 116074). This study was supported by grants to L. H. L.’s institution from the Swedish Research Council (grants 2013-23897-104604-23 and 523-2014-2336), the Swedish Heart Lung Foundation (grants 20150557 and 20170841), and the Stockholm County Council (grants 20140220 and 20170112). F. W. Asselbergs is supported by UCL Hospitals NIHR Biomedical Research Centre. I. Vaartjes is supported by the Dutch Heart Foundation, a part of ‘Facts and Figures’.

**Supporting information**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Table S1.** Missing data baseline characteristics and variables included in the multiple imputation for SwedeHF.

**Table S2.** Baseline characteristics of the external validation cohort (CHECK-HF) including missing data in percentages.

**Table S3.** Prediction thresholds for maximizing overall accuracy or sensitivity + specificity.

**Table S4.** External validation of the models in CHECK-HF.

**Table S5.** Simplified logistic model (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR) for EF cut-off ≥50%.

**Table S6.** Simplified logistic model (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR) for EF ≥40%.

**Table S7.** Sensitivity analysis of the logistic model EF ≥50% without NT-proBNP.

**Table S8.** Sensitivity analysis of the logistic model EF cut-off ≥40% without NT-proBNP.

**Table S9.** Sensitivity analysis of the logistic model EF ≥50% without NT-proBNP and NYHA class.

**Table S10.** Sensitivity analysis of the logistic model EF ≥40% without NT-proBNP and NYHA class.

**Table S11.** Simplified multinomial model (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR) predicting EF ≥50%.

**Table S12.** Sensitivity analysis of the multinomial model without NT-proBNP.

**Table S13.** Sensitivity analysis of the multinomial model without NT-proBNP and NYHA class.

**Figure S1.** (A) Study flow SwedeHF. (B) Study flow CHECK-HF.

**Figure S2.** Discrimination and calibration of the simplified logistic model (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR) predicting EF ≥50%.

**Figure S3.** Discrimination and calibration simplified logistic model (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR) predicting EF cut-off ≥40%.

**Figure S4.** Discrimination and calibration of the logistic model EF cut-off ≥50% without NT-proBNP.

**Figure S5.** Discrimination and calibration of the logistic model EF cut-off ≥40% without NT-proBNP.

**Figure S6.** Discrimination and calibration of the logistic model EF cut-off ≥50% without NT-proBNP and NYHA class.

**Figure S7.** Discrimination and calibration of the logistic model EF cut-off ≥40% without NT-proBNP and NYHA class.

**Figure S8.** Simplified multinomial model discrimination and calibration (i.e. not including NT-proBNP, NYHA class, mean arterial pressure, heart rate, BMI and eGFR).
Figure S9. Discrimination and calibration of the multinomial model without NT-proBNP.

Figure S10. Discrimination and calibration of the multinomial model without NT-proBNP and NYHA class.

References


