Family Practice, 2023, **40**, 188–194
https://doi.org/10.1093/fampra/cmac069
Advance access publication 2 July 2022
Diagnostic Studies

TARGET-HF: developing a model for detecting incident heart failure among symptomatic patients in general practice using routine health care data

Lukas De Clercq[1], Martijn C. Schut[2], Patrick M. M. Bossuyt[3], Henk C. P. M. van Weert[1], M. Louis Handoko[4], Ralf E. Harskamp[1,*]

[1]Department of General Practice, Amsterdam UMC, Location Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[2]Department of Medical Informatics, Amsterdam UMC, Location Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[3]Department of Public Health and Clinical Epidemiology, Amsterdam UMC, Location Academic Medical Center, University of Amsterdam, Amsterdam, The Netherlands
[4]Department of Cardiology, Amsterdam UMC, Location VU Medical Center, Vrije Universiteit, Amsterdam, The Netherlands

*Corresponding author: Department of General Practice, Amsterdam UMC, Amsterdam Public Health and Amsterdam Cardiovascular Sciences Research Institutes, University of Amsterdam, Room J2-126, Meibergdreef 9, 1105 AZ, Postbox 22660, 1100 DD Amsterdam, The Netherlands.
E-mail: r.e.harskamp@amsterdamumc.nl

Background: Timely diagnosis of heart failure (HF) is essential to optimize treatment opportunities that improve symptoms, quality of life, and survival. While most patients consult their general practitioner (GP) prior to HF, the early stages of HF may be difficult to identify. An integrated clinical support tool may aid in identifying patients at high risk of HF. We therefore constructed a prediction model using routine health care data.

Methods: Our study involved a dynamic cohort of patients (≥35 years) who consulted their GP with either dyspnoea and/or peripheral oedema within the Amsterdam metropolitan area from 2011 to 2020. The outcome of interest was incident HF, verified by an expert panel. We developed a regularized, cause-specific multivariable proportional hazards model (TARGET-HF). The model was evaluated with bootstrapping on an isolated validation set and compared to an existing model developed with hospital insurance data as well as patient age as a sole predictor.

Results: Data from 31,905 patients were included (40% male, median age 60 years) of whom 1,301 (4.1%) were diagnosed with HF over 124,676 person-years of follow-up. Data were allocated to a development (n = 25,524) and validation (n = 6,381) set. TARGET-HF attained a C-statistic of 0.853 (95% CI, 0.834 to 0.872) on the validation set, which proved to provide a better discrimination than C = 0.822 for age alone (95% CI, 0.801 to 0.842, $P < 0.001$) and C = 0.824 for the hospital-based model (95% CI, 0.802 to 0.843, $P < 0.001$).

Conclusion: The TARGET-HF model illustrates that routine consultation codes can be used to build a performant model to identify patients at risk for HF at the time of GP consultation.

Key words: decision support techniques, dyspnoea, early diagnosis, heart failure, oedema, primary health care

Background

Heart failure (HF) is a syndrome characterized by the heart's inability to meet the metabolic needs of the body. The underlying conditions are often multifactorial and may include comorbidities, such as coronary artery disease, hypertension, diabetes mellitus, or valvular disease.[1,2] Overall, around 2% of adults are diagnosed with HF, which increases to >10% over the age of 70 years.[1] Median life expectancy after diagnosis is 5 years, and the number of patients with HF is expected to double over the next decades. Timely diagnosis is important to allow the initiation of treatments that can improve outcomes, both in terms of mortality and quality of life. Leaders in the field of HF strongly propose to prioritize the aim of future efforts towards early detection and treatment of HF in order to alter the course of disease and limit further deterioration.[3,4]

The key to improve HF detection lies in the community, particularly when focussed on a population in which HF is at least moderately prevalent.[5] In this regard, a good starting population would therefore be patients who consult their general practitioner (GP) with symptoms associated with HF. Two hallmark symptoms of HF are shortness of breath (dyspnoea) and peripheral oedema; both are unfortunately non-specific in nature, more often than not arising from conditions other than HF.[6–8] Guidelines recommend that people presenting such symptoms to their GP get a natriuretic peptide test as well as an electrocardiogram, with referral for imaging and/or cardiologist review when either is abnormal.[1] Still, GPs selectively order these tests based on their perception of the patient's risk, resulting in selective diagnostic verification and diagnostic delay in patients not deemed at-risk. A report by the British Heart Foundation indicates that nearly 8 in 10 patients had visited their GP over the previous 5 years with symptoms associated with HF, but had not been diagnosed as such prior to emergency hospital admission.[9]

A user-friendly diagnostic support system should be developed to aid GPs in improving risk stratification. Unfortunately, existing HF risk prediction models are not fit for this task as

Key messages

- The key to improving heart failure (HF) detection lies in the community.
- The general practitioner (GP) has access to the community and its medical history.
- GP routine care data contain sufficient information for HF risk modelling.
- The developed TARGET-HF model does not require laboratory tests or other measurements.
- TARGET-HF outperforms existing secondary care models and age alone.
- The model is a promising candidate for an integrated clinical support tool.

they require actions to be performed, such as an ECG or laboratory tests, and were not developed for diagnostic support purposes.[5,10] Our primary objective was therefore to construct a model for incident HF based on known risk factors, yet only employing preregistered/routine care data available in the patients' electronic primary care health records. We compare this model with an existing insurance claims-based model[2] as well as with a model employing patient age as sole predictor.

Methods

We reported this study in accordance with the Transparent Reporting of a Multivariable Prediction Model for Individual Prognosis or Diagnosis statement.[11]

Data sourcing

Data were collected in a historical, dynamic cohort of patients registered at the GPs affiliated with the Academic Huisartsennetwerk Amsterdam, a non-profit organization in charge of region-based acquisition of GP data for research purposes. The dataset acquired in this network encompasses routine care data from more than 600,000 registered patients of 50 practices across sub-municipalities in the cities of Amsterdam and Almere. As GPs in the Netherlands function as gateway to the healthcare system, registration with a practice is *de facto* required, with estimated registration rates exceeding 99.9%.[12] Patients at participating practices can opt out of data sharing at any point in time, yet are noted by the network to do so rarely (<0.75%). In light of this information, we deem this dataset to accurately represent the community in the areas covered by the practices. Due to the nature of primary care, where patients can move to a practice outside the reach of the network, follow-up is not guaranteed across any time period.

The database is housed on a secure server of the Department of General Practice of the Amsterdam University Medical Center (AUMC). Researchers can perform analyses on pseudonymized subsets on a secure digital environment. Studies that leverage this are not classified as human subjects research by the AUMC Medical Ethics Committee provided they abide by the standardized isolation and deidentification procedures. Our dataset consists of patients' demographics, previous consultations, and recurring issues/chronic conditions dubbed "Episodes". Each consultation's reason (complaint, symptom, or condition) is coded using the International Classification of Primary Care (ICPC).[13]

Participants

Our historical cohort has an entry date of 2011 January 1 and an exit date of 2020 December 31, for a total observation period of 10 years. It includes all patients registered in the database at the time of extraction (2021 June 2) who met the following three eligibility criteria: (a) the patient has a consultation with an ICPC code for dyspnoea (K02, R02) or oedema (K07) occurring within the inclusion period at a point where the patient is 35 years or older (the "index" consultation) with (b) at least one consultation thereafter and (c) no registered HF diagnoses before the index consultation.

Outcome of interest

All potential HF diagnoses found in the patient records were verified manually under close scrutiny of a panel (LDC, RH) with expertise in general practice, cardiovascular medicine, and medical data processing. A search was initiated for identification of HF diagnoses, where episode records were evaluated for ICPC codes (K77, K84.03) and a series of textual searches of the GP's notes for terms indicating a HF diagnosis. The regular expressions used can be found in Supplementary material S1. Episodes with a HF code match were deemed a valid diagnosis unless the accompanying notes indicate otherwise (e.g. expression of doubt, differential diagnoses,…); those with only a textual match were considered invalid unless the context of said match proved otherwise.

Predictors

We searched the literature for population-based HF hazard prediction models to identify variables of interest. Based on the risk factors identified by two systematic reviews, reported by Yang et al.[5] and Sahle et al.,[10] and the availability thereof as ICPC codes, we arrive at a set of 2 demographic values—sex and age—and 14 medical history variables: tobacco use, alcohol abuse, obesity, material deprivation (poverty/financial issues), family history of cardiovascular disease (CVD), hypertension, diabetes mellitus, coronary artery disease, atrial fibrillation, heart murmur, valvular heart disease, stroke, chronic obstructive pulmonary disease, and chronic kidney disease. This entire procedure is detailed in Supplementary material S2.

All conditions present prior to each patient's index consultation were considered antecedent conditions. The presence thereof was established through the means of an ICPC code search, yielding a structured medical history suitable for our regression models. All relevant ICPC codes are detailed in Supplementary Table S3.

Missing data

The variables regarding medical history represent the registration of conditions by the GP prior to the index consultation, not the actual presence of said condition within a patient. In that sense, they can—per definition—not be missing. Demographic values are entered for each patient as part of

their registration with a GP practice and were found to be complete for all included patients.

Statistical analysis methods

We calculated the incidence of first registered HF diagnoses in our cohort as events per 1,000 person-years, which we report for men and women separately and over different age ranges identical to those used by Goyal et al. in their analysis.[2] Patients were censored from incidence calculations for the time period beyond their HF diagnosis, when present. Incidence rate ratios between sexes are reported across the age groups with their exact Poisson confidence intervals.

A 20% outcome-stratified validation set of our cohort was isolated before the model development process to validate the acquired model. A cause-specific Cox proportional hazards model was chosen for its ability to handle right-censored data. Taking heed of the concerns Sahle et al. had regarding previous incident HF risk prediction models, we attempted to avoid ill-advised model development practices such as stepwise variable selection or categorization of continuous variables.[10]

As in Goyal et al., we modelled hazard ratios under three conditions: independently, adjusted for sex and age, and adjusted for all included variables.[2] This exploratory phase was followed by using the full set of variables to generate a final predictive model using an L_1-penalty for regularization and feature selection (LASSO). This technique shrinks the model's coefficients, the rate of which is determined by a penalty parameter, potentially eliminating them from the model altogether as they shrink to zero.[14] The optimal penalty was selected through a randomized search applied to a 5-fold cross-validation, after which the final model (TARGET-HF) was trained on the entire development set with the acquired penalty value. One alternative regularization technique—$L_{1/2}$ regularization—was evaluated but yielded no appreciable improvements in calibration. More information regarding this comparison and the results thereof can be found in Supplementary material S4, Supplementary Table S7 and Supplementary Figure S8. All models were constructed and evaluated in Python using the Lifelines[15] and Scikit-survival[16] libraries.

Predictive performance of the model was compared to a baseline of age as sole predictor as well as the "outpatient" model developed by Goyal et al. based on hospital insurance data. The latter was chosen for its use of routine care variables, i.e. not requiring additional measurements or tests, all of which were available as ICPC codes in our dataset. All models were evaluated using Harrell's C-statistic[17] on the validation set as a whole as well as stratified by age groups identical to those used by Goyal et al. Following the principle of Heagerty and Zheng[18] and the weighted method of Uno et al.,[19] we calculated area under the cumulative/dynamic receiver–operator curve ($AUROC^{C,D}$) over the first 5 years of follow-up in 1-month increments. Confidence intervals were calculated through bootstrapping (1,000 iterations).

Results

Cohort analysis

A total of 31,905 patients met our cohort criteria, a flowchart of which can be found in Fig. 1. Our cohort was 40% male and the median age was 60 years; further baseline characteristics are summarized in Table 1.

Heart failure incidence

Inspection of the episodes yielded a total of 4,731 ICPC code matches and 1,033 additional textual matches, with 698 (8.3%) and 393 (29.5%) of those, respectively, identified as false positive. The textual searches found an additional 10.9% of patients with a HF diagnosis when compared to searching on ICPC codes alone, as can be seen in Fig. 1. After removal of the 2,129 patients with diagnosed HF prior to their index consultation, we were left with 1,301 patients (4.08%) diagnosed with HF over 124,676 person-years of follow-up in our cohort.

We observed a HF incidence rate of 10.44 per 1,000 person-years (95% CI, 9.88 to 11.02), with 12.96 per 1,000 person-years for men (95% CI, 11.95 to 14.04) and 8.94 (95% CI, 8.29 to 9.63) and women. A full overview of rates and rate ratios across sexes and age ranges can be found in Supplementary Table S5.

Predictors

Our development set for training the models held data from 25,524 persons, of which 1,041 (4.1%) had a registered incident HF event in their follow-up period. With a total of 16 variables, this puts our models (prior to feature selection) at over 65 events per variable. The hazard coefficients of the unregularized models and their corresponding *P*-values are reported in Supplementary Table S6.

The L_1-regularized model, TARGET-HF, saw the coefficients of two of its variables—tobacco use and family history of CVD—reduced to zero, effectively eliminating them from the model. Its hazard ratios, in comparison against those from Goyal et al.'s outpatient model, are shown in Fig. 2.

Predictive performance

Our validation set held data from 6,381 persons, of which 260 (4.07%) carried an incident HF event during their follow-up period. For this set, Harrel's C-statistic was 0.853 for TARGET-HF (95% CI, 0.834 to 0.872), which proved to outperform C = 0.824 for Goyal et al.'s model (95% CI, 0.802 to 0.843, $P < 0.001$) and C = 0.822 for the baseline model with age only (95% CI, 0.801 to 0.842, $P < 0.001$). Further C-statistics of the validation set stratified by age can be found in Table 2. Classification performances and their confidence intervals across time in the form of the $AUROC^{C,D}$ can be observed in Table 3 and Fig. 3.

The C-statistic within the development set was 0.812 for TARGET-HF (95% CI, 0.799 to 0.826), outperforming C = 0.788 for Goyal et al.'s model (95% CI, 0.775 to 0.803, $P < 0.001$) and C = 0.784 for a baseline model based on age alone (95% CI, 0.770 to 0.798, $P < 0.001$).

Discussion

Timely detection of HF starts with accurate identification of patients who should undergo further diagnostic work-up. On this premise, we developed the TARGET-HF model using routine primary care data. We demonstrated that it is feasible to construct a prediction model that outperforms both age as a sole predictor and an existing community-based prediction model. Moreover, this discriminatory advantage remained consistent when the model was applied to an isolated validation set.

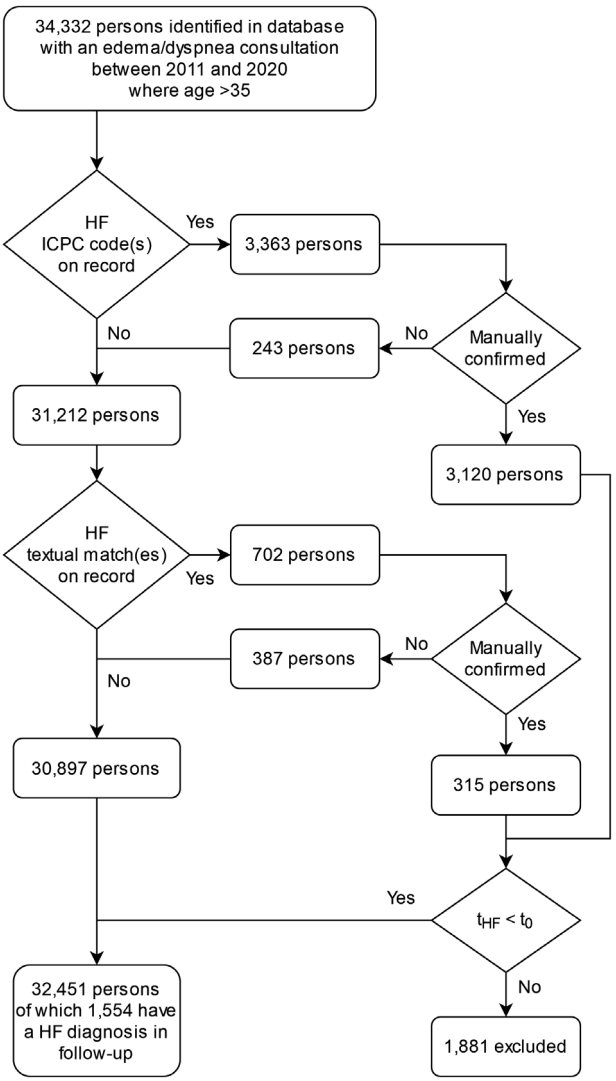

Fig. 1. Flowchart depicting in- and exclusion criteria for the cohort used in the development and evaluation of the TARGET-HF model. t_{HF} signifies the time of the identified heart failure (HF) diagnosis registration, t_0 symbolizes the reference consultation.

Strengths and limitations

Our study has a number of strengths. The advantage of using routine care data is that it involves an unselected and therefore representative sample of a population in a metropolitan area. Moreover, the data are rich in both structured and unstructured content, such as consultation notes and abstracts of specialist letters, which allowed us to more accurately identify HF cases.

Working with routine care data also comes with challenges. First, our methods for identifying HF in the follow-up were rigorous but not fool-proof, the large ratio of discarded episodes based on their content gave rise to suspicion towards episodes lacking elaborate descriptions. Few opportunities remain to remediate this situation aside from external linkage with hospital registrations, though this is hampered by privacy concerns and the lack of a centralized medical record or unique identifier. Second, our predictors are proxies for incidence of the conditions they represent, meaning that we did not take into account the time since a code was registered. There may be performance gains in using algorithms capable of modelling this temporal information. A related limitation is the lack of correction for patients' differing inclinations and/or motives to visit their GP and/or report symptoms or lifestyle complaints. Feasible variables to include to account for this are various derivations of consultation count/frequency prior to the index consultation.

Selective reporting may also be an issue. Our predictors—the registration of a code at any point in the past—are used as such and may or may not reflect the presence or absence of the conditions they represent. In that sense, they ought to be considered proxies. Especially for lifestyle risk factors, there is a concern for underreporting and reporting bias, as they will only be registered when the patient brings these subjects up or when the GP chooses to register a lifestyle-related code for a consultation for a related complaint. This is exemplified by lifetime tobacco use, which in our cohort is only around 11% for both men and women, whereas a large-scale questionnaire established a far larger proportion in a population sample.[20] Nonetheless, these proxies appear to be able to function as predictors for incident HF, as shown by our models.

Prior studies

To the best of our knowledge, this study is the first to build a prediction model on HF using routine primary care data. However, community-based models do exist that predict HF. A systematic review of these models found that the strongest associations with incident HF have been observed for age, coronary artery disease, diabetes mellitus, hypertension, and smoking.[5] When comparing these risk factors with the findings from our analysis we observe similar associations, especially for the demographic values. There were also notable differences, for instance, our model's low weight on hypertension, a factor that typically contributes fairly heavily in other incident HF models.[5] Another interesting observation is that reported material deprivation yields a coefficient indicating a lower risk compared to the baseline hazard, contradicting prior findings indicating that material deprivation should be viewed as a risk factor for HF.[21] While speculative, we postulate that this may be related to care avoidance and subsequent delayed diagnoses, as well as fewer diagnostic procedures and/or poor registration of incident HF in those with recorded material deprivation versus those without.

Clinical implications

HF is a major health problem of increasing prevalence that severely affects the quality of life and shortens lives.[22] It is often not diagnosed early enough to take full advantage of ameliorating medication.[23] A reliable tool that can help detect high-risk patients for HF in the community would be an important asset to tackle this. Using TARGET-HF, we may provide GPs with an important first building block to improve early detection of HF. The model could easily be integrated in existing primary care electronic health record systems, where it could run in the background, to be activated when a patient is evaluated with symptoms suggestive for HF. It could perhaps work in tandem with the recent BEAT-HF campaign, which aims to increase clinicians' awareness of anyone who presents with breathlessness, exhaustion, and ankle swelling.[4]

Future studies

The prototype that we have developed appears promising, yet further validation is warranted prior to implementation.

Table 1. Baseline characteristics of patients reporting with dyspnoea and/or ankle oedema between 2011 and 2020, with and without heart failure (HF) in their follow-up.

	HF (n = 1,301)	No HF (n = 30,604)	P-value
Age[a] (median, 25–75th)	77 (67–84)	59 (49–71)	<0.001
Years of recorded history[a] (median, 25–75th)	2.2 (0.8–4.0)	3.6 (1.5–6.0)	<0.001
Years of follow-up[a] (median, 25–75th)	4.2 (2.2–6.2)	3.6 (1.5–6.1)	<0.001
Years within GP network (median, 25–75th)	7.1 (5.5–9.6)	9.0 (6.5–9.7)	<0.001
Male sex	601 (46.2%)	12,047 (39.4%)	<0.001
Tobacco use	101 (7.8%)	3,320 (10.9%)	<0.001
Alcohol abuse	66 (5.1%)	1,231 (4.0%)	0.062
Obesity	141 (10.8%)	3,950 (12.9%)	0.028
Material deprivation	8 (0.6%)	445 (1.5%)	0.008
Family history of cardiovascular disease	3 (0.2%)	162 (0.5%)	0.167
Hypertension	735 (56.5%)	10,650 (34.8%)	<0.001
Diabetes mellitus	433 (33.3%)	5,170 (16.9%)	<0.001
Coronary artery disease	404 (31.1%)	3,507 (11.5%)	<0.001
Atrial fibrillation	288 (22.1%)	1,598 (5.2%)	<0.001
Heart murmur	17 (1.3%)	261 (0.9%)	0.092
Valvular heart disease	167 (12.8%)	1,025 (3.4%)	<0.001
Stroke	144 (11.1%)	1,431 (4.7%)	<0.001
Chronic obstructive pulmonary disease	256 (19.7%)	2,806 (9.2%)	<0.001
Chronic kidney disease	229 (17.6%)	1,911 (6.2%)	<0.001

Distributional differences are evaluated with Fisher's exact test for proportions and a Mann–Whitney U test for durations.
[a]Durations established relative to the index consultation.

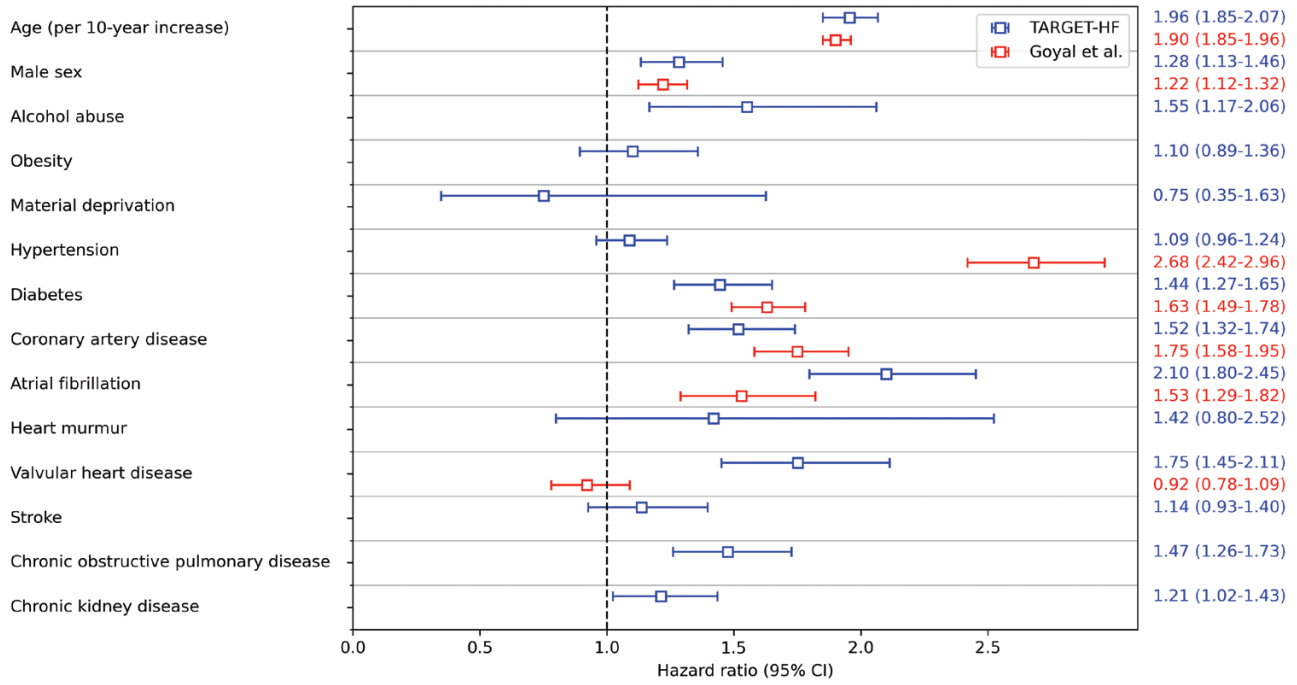

Fig. 2. Hazard ratio comparison between the TARGET-HF model and Goyal et al.'s outpatient model.

Moreover, interventional studies would then be required to evaluate whether integrating this algorithm in GP practice does indeed result in improved HF detection, initiation of adequate treatment, and ultimately improved clinical outcomes. However, prior to these steps, we propose to evaluate whether additional techniques, such as natural language processing, could further improve the prediction of HF or diagnostic coding thereof. Textual queries provided us with an additional 11% of HF diagnoses not caught by ICPC codes, a non-negligible number for such a serious heart condition. In

Table 2. Harrel's C-statistic of predicting incident heart failure on the validation set of a cohort of patients reporting with dyspnoea and/or ankle oedema between 2011 and 2020, evaluated for TARGET-HF and compared to Goyal et al.'s outpatient model and age as a sole predictor.

C-statistic	Baseline (age)	TARGET-HF	Goyal et al. (outpatient)
All ages (n = 6,381)	0.822 (0.801–0.842)	**0.853** (0.834–0.872)	0.824 (0.803–0.843)
35–54	0.721 (0.510–0.857)	0.831 (0.627–0.943)	**0.833** (0.778–0.888)
55–64	0.538 (0.438–0.645)	**0.718** (0.606–0.814)	0.650 (0.551–0.739)
65–74	0.638 (0.572–0.693)	**0.732** (0.662–0.792)	0.631 (0.568–0.693)
75+	0.632 (0.581–0.678)	**0.688** (0.576–0.678)	0.626 (0.581–0.678)

The values in bold represent the model with the best performance per age category.

Table 3. Area under the cumulative/dynamic receiver–operator curve ($AUROC^{C,D}$) of predicting incident heart failure on the validation set of a cohort of patients reporting with dyspnoea and/or ankle oedema between 2011 and 2020, evaluated at several points in time after the index consultation for TARGET-HF and compared to Goyal et al.'s outpatient model and age as a sole predictor.

$AUROC^{C,D}$	Baseline (age)	TARGET-HF	Goyal et al. (outpatient)
1 year	0.822 (0.791–0.852)	**0.855** (0.826–0.882)	0.825 (0.793–0.854)
2 years	0.849 (0.824–0.873)	**0.878** (0.855–0.899)	0.841 (0.814–0.864)
3 years	0.850 (0.826–0.874)	**0.888** (0.867–0.908)	0.853 (0.829–0.875)
4 years	0.854 (0.827–0.878)	**0.885** (0.860–0.908)	0.855 (0.830–0.878)
5 years	0.866 (0.841–0.889)	**0.894** (0.871–0.914)	0.861 (0.837–0.889)

The values in bold represent the model with the best performance per point in time.

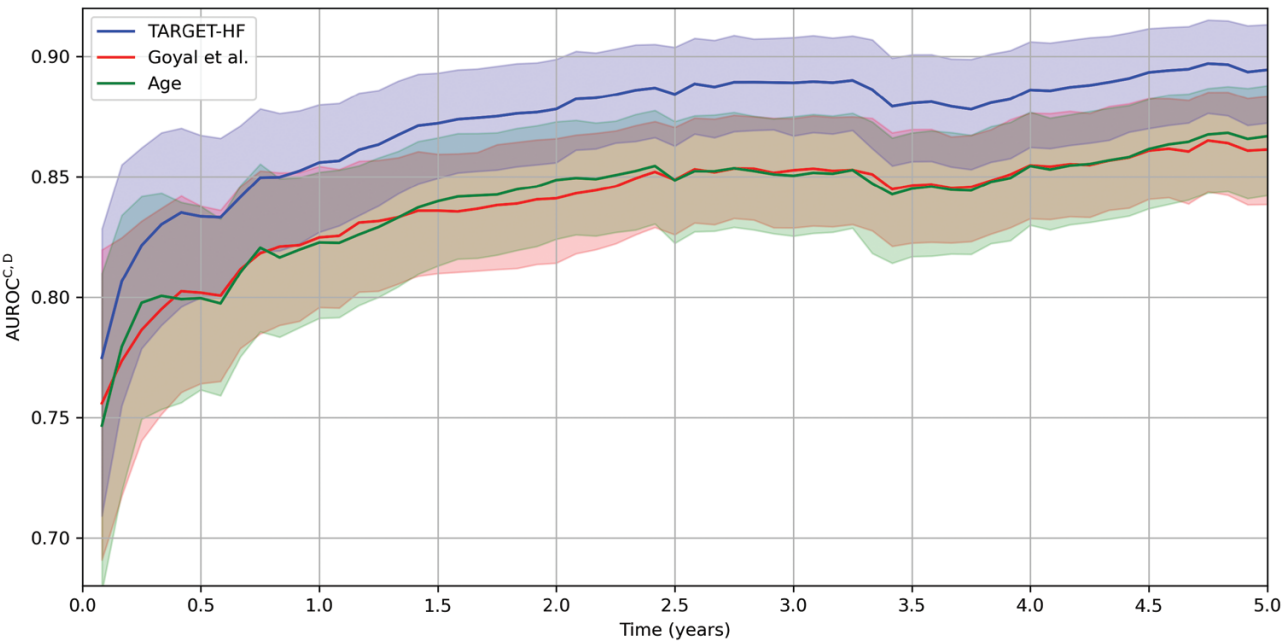

Fig. 3. Area under the cumulative/dynamic receiver–operator curve ($AUROC^{C,D}$) of TARGET-HF compared to Goyal et al.'s outpatient model and age as a sole predictor in predicting incident heart failure, calculated on a 1-month resolution.

other words, there is still work to be done and we believe that there may be predictive potential in the unstructured data of GP patient files that has thus far been untapped.

Conclusion

The TARGET-HF model illustrates that consultation codes found in routine primary care data can be used to build an effective predictive model to identify patients at risk for HF at the time of GP consultation. Moreover, we found that the model outperformed both age as a sole predictor and an existing community-based prediction model based on hospital insurance data.

Supplementary material

Supplementary material is available at *Family Practice* online.

Funding

This work was supported by departmental resources and a seeding grant from the Amsterdam Public Health Research Institute (TARGET-HF).

Ethical approval

This study was exempted from formal ethics review, the data protection impact assessment was evaluated by the data protection officer of the Amsterdam UMC-AMC location on 2020 March 26 (dPIA-target-hf-27-2-2020).

Conflict of interest

None.

Data availability

All relevant data produced in the present work are contained in the manuscript. Requests for access to original data files can be directed at the Amsterdam UMC general practice network (https:/www.amsterdamumc.org).

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, Burri H, Butler J, Čelutkienė J, Chioncel O, et al. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: developed by the Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC) With the special contribution of the Heart Failure Association (HFA) of the ESC. *Eur Heart J*. 2021;42(36):3599–3726.
2. Goyal A, Norton CR, Thomas TN, Davis RL, Butler J, Ashok V, Zhao L, Vaccarino V, Wilson PWF. Predictors of incident heart failure in a large insured population: a one million person-year follow-up study. *Circ Heart Fail*. 2010;3(6):698–705.
3. Taylor CJ, Ordóñez-Mena JM, Roalfe AK, Lay-Flurrie S, Jones NR, Marshall T, Hobbs FDR. Trends in survival after a diagnosis of heart failure in the United Kingdom 2000–2017: population based cohort study. *BMJ*. 2019;364:l223.
4. Taylor CJ, Hartshorne-Evans N, Satchithananda D, Hobbs FR. FASTer diagnosis: time to BEAT heart failure. *BJGP Open*. 2021;5(3):BJGPO.2021.0006. doi:10.3399/BJGPO.2021.0006.
5. Yang H, Negishi K, Otahal P, Marwick TH. Clinical prediction of incident heart failure risk: a systematic review and meta-analysis. *Open Heart*. 2015;2(1):e000222.
6. Ely JW, Osheroff JA, Chambliss ML, Ebell MH. Approach to leg edema of unclear etiology. *J Am Board Family Med*. 2006;19(2):148–160.
7. Cho S, Atwood JE. Peripheral edema. *Am J Med*. 2002;113(7):580–586.
8. Frese T, Sobeck C, Herrmann K, Sandholzer H. Dyspnea as the reason for encounter in general practice. *J Clin Med Res*. 2011;3(5):239–246.
9. Blake I. Heart failure hospital admissions rise by a third in five years; 2019 [accessed 2021 Dec 5]. https://www.bhf.org.uk/what-we-do/news-from-the-bhf/news-archive/2019/november/heart-failure-hospital-admissions-rise-by-a-third-in-five-years.
10. Sahle BW, Owen AJ, Chin KL, Reid CM. Risk prediction models for incident heart failure: a systematic review of methodology and model performance. *J Card Fail*. 2017;23(9):680–687.
11. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ*. 2015;350:g7594.
12. Poortvliet MC, Lamkaddem M, Devillé W. *Niet op naam ingeschreven (NONI) bij de huisarts: Inventarisatie en gevolgen voor de ziekenfondsverzekerden*. Utrecht (the Netherlands): NIVEL; 2005.
13. Lamberts H, Wood M. *ICPC: International Classification of Primary Care*. Oxford (UK): Oxford University Press; 1987.
14. Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med*. 1997;16(4):385–395.
15. Davidson-Pilon C. *Lifelines, survival analysis in Python (v0.26.3)*. Geneva, Switzerland: Zenodo; 2021.
16. Pölsterl S. Scikit-survival: a library for time-to-event analysis built on top of scikit-learn. *J Mach Learn Res*. 2020;21(212):1–6.
17. Harrell FE, Califf RM, Pryor DB, Lee KL, Rosati RA. Evaluating the yield of medical tests. *JAMA*. 1982;247(18):2543–2546.
18. Heagerty PJ, Zheng Y. Survival model predictive accuracy and ROC curves. *Biometrics*. 2005;61(1):92–105.
19. Uno H, Cai T, Tian L, Wei L-J. Evaluating prediction rules for *t*-year survivors with censored regression models. *J Am Stat Assoc*. 2007;102(478):527–537.
20. Degenhardt L, Chiu W-T, Sampson N, Kessler RC, Anthony JC, Angermeyer M, Bruffaerts R, de Girolamo G, Gureje O, Huang Y, et al. Toward a global view of alcohol, tobacco, cannabis, and cocaine use: findings from the WHO World Mental Health Surveys. *PLoS Med*. 2008;5(7):e141.
21. Hippisley-Cox J, Coupland C. Development and validation of risk prediction equations to estimate future risk of heart failure in patients with diabetes: a prospective cohort study. *BMJ Open*. 2015;5(9):e008503.
22. Conrad N, Judge A, Tran J, Mohseni H, Hedgecott D, Crespillo AP, Allison M, Hemingway H, Cleland JG, McMurray JJV, et al. Temporal trends and patterns in heart failure incidence: a population-based study of 4 million individuals. *Lancet*. 2018;391(10120):572–580.
23. Boffa U, McGrady M, Reid CM, Shiel L, Wolfe R, Liew D, Campbell DJ, Stewart S, Krum H, et al. SCReening Evaluation of the evolution of new heart failure study (SCREEN-HF): early detection of chronic heart failure in the workplace. *Aust Health Rev*. 2017;41(2):121–126.