BMJ Open

Development and validation of a risk stratification model for prediction of disability and hospitalisation in patients with heart failure: a study protocol

Luis García-Olmos,[1,2] Francisco Rodríguez-Salvanés,[3] Maurice Batlle-Pérez,[4] Río Aguilar-Torres,[5] Carlos Porro-Fernández,[4] Alfredo García-Cabello,[4] Montserrat Carmona,[2,6] Sergio Ruiz-Alonso,[7] Sofía Garrido-Elustondo,[1,2] Ángel Alberquilla,[2,8] Luis María Sánchez-Gómez,[2,9] Ricardo Sánchez de Madariaga,[6] Elena Monge-Navarrete,[10] Luis Benito-Ortiz,[11] Nicolás Baños-Pérez,[10] Amaya Simón-Puerta,[12] Ana Belén López Rodríguez,[13] Miguel Ángel Martínez-Álvarez,[12] María Ángeles Velilla-Celma,[12] María Isabel Bernal-Gómez,[11] CHIC Group

To cite: García-Olmos L, Rodríguez-Salvanés F, Batlle-Pérez M, *et al*. Development and validation of a risk stratification model for prediction of disability and hospitalisation in patients with heart failure: a study protocol. *BMJ Open* 2017;**7**:e014840. doi:10.1136/bmjopen-2016-014840

Received 24 October 2016
Revised 10 April 2017
Accepted 11 April 2017

For numbered affiliations see end of article.

Correspondence to
Dr Luis García-Olmos; lgarciao@salud.madrid.org

ABSTRACT

Background Chronic heart failure (CHF) reduces quality of life and causes hospitalisation and death. Identifying predictive factors of such events may help change the natural history of this condition.

Aim To develop and validate a stratification system for classifying patients with CHF, according to their degree of disability and need for hospitalisation due to any unscheduled cause, over a period of 1 year.

Methods and analysis Prospective, concurrent, cohort-type study in two towns in the Madrid autonomous region having a combined population of 1 32 851. The study will include patients aged over 18 years who meet the following diagnostic criteria: symptoms and typical signs of CHF (Framingham criteria) and left ventricular ejection fraction (EF)<50% or structural cardiac lesion and/or diastolic dysfunction in the presence of preserved EF (EF>50%). Outcome variables will be(a) Disability, as measured by the WHO Disability Assessment Schedule V.2.0 Questionnaire, and (b) unscheduled hospitalisations. The estimated sample size is 557 patients, 371 for predictive model development (development cohort) and 186 for validation purposes (validation cohort). Predictive models of disability or hospitalisation will be constructed using logistic regression techniques. The resulting model(s) will be validated by estimating the probability of outcomes of interest for each individual included in the validation cohort.

Ethics and dissemination The study protocol has been approved by the Clinical Research Ethics Committee of La Princesa University Teaching Hospital (PI-705). All results will be published in a peer-reviewed journal and shared with the medical community at conferences and scientific meetings.

Strengths and limitations of this study

- Study based on a population attended in primary care, which is representative of the general population.
- Disability and predictive factors in patients with heart failure are poorly understood.
- As most admissions will take place at a single hospital, hospital-related influence on hospital admissions cannot be estimated.
- There is a possibility of patient loss at advanced disease stages.

BACKGROUND

Heart failure (HF) affects 1%–2% of the population aged over 40 years in Spain,[1] with prevalence progressively increasing due to population ageing. It is the leading cause of hospital admissions among the population aged over 65 years[2] and accounts for 5% of all hospitalisations in this age group. Following hospital admission, up to 50% of patients are readmitted after 6 months.[3] HF ranks among the processes known to have the highest impact on quality of life.[4 5] In Spain, HF caused 15% of cardiovascular deaths in 2009: average survival 5 years after diagnosis is 50%.[6] Patients suffering from HFnotonly present with high comorbidity but they also consume a high amount of healthcare resources.[7 8] HF accounts for 1%–2% of overall healthcare costs, with 60%–74% of such HF-related cost being attributable to hospitalisations.[9]

At present, effective treatments are available. Despite the fact that these treatments have achieved a reduction in mortality, long-term prognosis of such patients is not optimal. HF is an important cause of avoidable hospitalisations and, since hospitalisations represent the

greatest part of the healthcare costs generated by these patients, predictive models that could prevent hospitalisations have been sought for many years.

Survival for patients with HF is poorly predicted by physicians,[10] as is probably the risk of hospital readmissions.[11] From a clinical point of view, knowing the prognosis of patients with HF and the factors determining their conditionenables physicians and patients to take decisions and may contribute to reduce the gap in the translation of research results into clinical practice. Although research regarding predictive factors of hospital readmission has been intensified, the studies performed have been flawed by methodological problems and this has, in turn, generated confusion about the validity of the results underlying the need for publication of guidelines with recommendations to researchers.[12]

When it comes to health services, strategies for stratifying patients into different risk groups have been implemented, in order to help organise services by adjusting their type and intensity according to patients' needs. Models have been developed to stratify patients according to the risk of presenting with undesirable events, which are relevant to the health system, avoidable through healthcare interventions and predictable with the aid of existing data. Frequently, the development of such models has been specifically redirected to the task of predicting hospital readmissions.

From a community perspective, in a sample with more than 1000 incident HF cases monitored for an average of 4.7 years, the risk of hospitalisation was 86.6% per person-year, with more than half being readmitted on three or more occasions.[13] The cause of hospitalisation was HF in only 16.5% of cases versus non-cardiovascular reasons in 62% of cases.

In a systematic review, which included 30 studies that evaluated 26 different prediction models, four studies developed predictive models for patients with HF: allmodels displayed a low predictive capacity with a c-statistic of 0.6–0.7.[14] Consequently, their impact on prediction of undesirable events was low. Patients' own variables yielded inconsistent results, with differences depending on the population on which the study was conducted.

With regard to risk-stratification models in patients with HF, in 2008, a systematic review included 117 studies, most of which were based on administrative data and clinical records.[15] Five of these studies sought to develop statistical models capable of predicting rehospitalisation risks. Sociodemographic variables were analysed, as were comorbid conditions, HF severity indices and markers such as blood urea nitrogen (BUN), sodium, B-type natriuretic peptide (BNP), haemoglobin and troponin. Results were inconsistent, with low discriminative power, and no acceptable risk-stratification model was found.

A recent systematic review of models designed to predict hospitalisations, death or both among patients suffering from HF included 48 studies and 64 models. This review highlighted both the heterogeneity of the studies and the fact that the models were better at predicting death than hospitalisations.[16] In the models constructed, biological parameters prevailed and included: demographic variables, that is, age, sex and race; clinical variables such as comorbidities, cancer, diabetes, chronic obstructive pulmonary disease (COPD), renal failure, blood pressure, body mass index, the New York Heart Association (NYHA) functional classification and left ventricular ejection fraction (LVEF); biomarkers, for example, creatinine, BUN, sodium and BNP; andtype of medications. In general, few clinical variables with predictive capacity were found, and the results proved largely inconsistent.

Previous hospitalisations, number of hospital admissions, comorbidity and polypharmacy have been identified as predictive factors of rehospitalisations.[17 18] Introduction of variables reflecting socioeconomic level, social aspects of individuals or quality of life[19] have improved the discriminative power of the models. Shih *et al*[20] found that patients' functional capacity was better at predicting rehospitalisations than comorbidities in medically complex patients. Socioeconomic data are not usually included in health records but imputation to patient of the average hospitalisations of theircensus area has been used as a surrogate indicator.[21 22]

Currently, HF predictive models focus on the hospital setting and the prevention of rehospitalisations. These kinds of studies, based on hospital data, solely predict the risk of new events in a hospital population. In some cases, mortality predictors have been sought but we were unable to find any studies aimed at predicting the progress of disability. However, identification of individuals at risk who may not have been previously hospitalised affords an opportunity for wider interventions and may well allow for earlier actions targeted at preventing greater damage.

It is not clear whether it is better to develop local, regional or national models.[23] Models aimed at specific conditions have greater discriminative capacity than do models aimed at the whole population.[24]

Profiles of patients suffering from HF have changed, in that they are currently older, present with more comorbidities, tend to be polymedicated and have a high degree of disability.[25] Patients suffering from HF and a high degree of disability are readmitted and die more frequently;[26] consequently, identification of patients with disability will make it possible to develop specific preventive interventions[27] purpose designedto prevent such events. The existence of transculturally validated tools, such as the WHO Disability Assessment Schedule V.2.0 (WHODAS 2.0 questionnaire), enables this problem to be tackled.[28] The WHODAS 2.0 is an instrument designed to measure disability from a biopsychosocial perspective. It has been tested in 16 languages in 14 countries and has proved to have adequate metric properties for patients with chronic illness in Spain. This instrument enables the study of the interaction between a given patient's state of health and the characteristics of the environment in which he/she evolves.

Accordingly, we used theWHODAS V.2.0 to study the functioning capacity of patients with HF, cerebrovascular

illness and COPD,[29] and changes in such functioning capacity during the course of 30 months.[30] We confirmed that HF gave rise to more disability than did the other two conditions studied, affecting overall activities of daily living and participation in society. One-third of the patients presented with severe disability, which was more marked among women engaged in domestic activities of daily living and mobility. Patients with HF evolved towards a progressive deterioration in their disability, which was conditioned by their baseline situation, such that a one-point change in the baseline situation determined a 12% increase in the risk of developing severe/complete disability or dying.

HF serves as a paradigm of chronic illness, with a gradual increase in the need for healthcare and the intervention from professionals of different healthcare and social levels. A predictive model of undesirable events among patients with HF, developed from primary care and allowed for patient stratification according to their health service needs, would be useful in terms of facilitating personalised attention and yielding improved outcomes.

AIMS

Main objective

To develop and validate a stratification model for classifying patients suffering from chronic heart failure, according to their degree of disability and risk of hospitalisation during the year following their inclusion in the study.

Secondary objectives

1. To describe the sociodemographic and clinical characteristics, and comorbidity of a population-based (ie,health card based) cohort of patients suffering from HF.
2. To identify which characteristics of the patients included in the cohort determine degree of disability after 1 year of follow-up.
3. To identify which characteristics of the patients included in the cohort determine unscheduled hospitalisations, for any reason, after 1 year of follow-up.
4. To identify which characteristics of the patients included in the cohort determine death due to any cause, after 1 year of follow-up.

METHODS AND ANALYSIS

Design

Concurrent, prospective, cohort-type study.

Participants

Population aged over 18 years attended at five healthcare centres situated in two towns in the Madrid autonomous region (*Comunidad de Madrid*), where the study is to be carried out. The two towns have a total combined population of 132 851, 46% aged over 40 years and 10% aged over 65 years.

Patient recruitment

Patients will be identified in three different ways:

1. from primary care electronic health records (EHRs): all patients diagnosed with HF, International Classification of Primary Care code K77, will be retrieved;
2. from the hospital minimum basic data set (MBDS): the study population has been served by a single referral hospital for the last 10 years. Until 2007, however, there was a different referral hospital. To avoid any possible loss of patients and data in a case where some patient might still be undergoing treatment at the former referral hospital, we have, therefore, decided to identify and retrieve all patients who are shown in the MBDS of the two hospitals as having HF (*International Classification of Diseases, ninth revision, Clinical Modification (ICD-9-CM)* codes 428.x and 402.01, 402.11 and 402.91) as their principal or secondary diagnosis, and who reside in the study area;
3. from consultations with general practitioners (GPs) at the healthcare centres participating in the study.

Lists with patients corresponding to each healthcare centre will be generated by cross referencing data against the respective health card code numbers.

Inclusion criteria

Any patient diagnosedwith HF, who meets the following diagnostic criteria (case definition):

1. symptoms and typical signs of HF (Framingham criteria): (1) major criteria : night paroxystic dyspnoea, orthopnoea, high jugular pressure, crackles, third tone, cardiomegaly, interstitial lung oedema in chest radiography, 4.5 kg weight loss with treatment for HF, and (2) minor criteria: bilateral oedema, night cough, dyspnoea with ordinary exertion, hepatomegaly, pleural effusion and tachycardia. Presence of two major criteria or presence of one major and two minor criteria are suggestive of HF; and
2. LVEF (<50%, Simpson's integration method) that includes either subjects with reduced or mid-range LVEF according to the European Society of Cardiology guidelines 2016[31] or relevant structural cardiac lesion and/or diastolic dysfunction (filling alteration grade two or higher, with E/e′ ratio >15 or elevated natriuretic peptide (NP) levels (BNP >35 pg/mL and/or N-terminal pro-B-type natriuretic peptide (NT-proBNP) >1295 pg/mL), in the absence of thromboembolic illness, pulmonary heart disease (cor pulmonale) or other causes of elevated proBNP). Similarly, we also plan to include any patient who underwent a recent echocardiography cardiac function study (<6 months before inclusion) and fulfils the above criteria;
3. in all cases, patients will be asked to give their informed consent.

Exclusion criteria

The following will be excluded from the study:

1. institutionalised patients,
2. patients suffering a terminal illness other than HF,
3. patients with a life expectancy of <6 months and
4. patients who refuse to participate in the study.

Determinations and variables

1. Dependent variables
   - Unscheduled hospitalisations obtained from the hospital MBDS and primary care EHRs;
   - WHODAS 2.0 scale scores: the 36-item version of the WHODAS 2.0 is a scale that is modelled on the International Classification of Functioning, Disability and Health (ICF) and measures disability. The final score is expressed as a percentage of the maximum theoretical score, with WHODAS 2.0 scores being interpreted as follows: 0%–4%, no disability; 5%–24%, low disability; 25%–49%, moderate disability; 50%–95%, severe disability and 96%–100%, complete or extreme disability. The questionnaire groups items into six domains, namely, understanding and communication, mobility, personal care, interpersonal relationships, everyday activities and labour activity and social participation. The WHODAS 2.0 (36-item version) has been widely tested and been shown to be an instrument that is both robust and easy to administer;
   - all-cause mortality obtained from death registries.
2. Independent variables
   - age, sex, marital status, work status, educational level, socioeconomic level (average income in subject's census area), echocardiogram with LVEF, NT-proBNP, drug use (angiotensin-converting enzyme inhibitors, angiotensin-II receptor blockers, beta blockers and diuretics), medical visits to the GP per year, medical visits to a nurse per year and previous hospitalisations;
   - comorbidity: this will be described using expanded diagnosis clusters (EDCs) based on the Johns Hopkins University Adjusted Clinical Group patient classification system,[32] which performs different groupings. On the basis of patients' diagnoses, this system groups diseases, diagnosis and health conditions having similar clinical characteristics into EDCs. There are 269 EDCs and each patient may have several of them. The system identifies 127 chronic EDCs that are expected to persist for 12 months or more.

We will study the presence of 127 chronic EDCs and the number of such EDCs in each patient. The number of EDCs and some EDCs that have been found to be relevant in other studies will be included as independent variables in the modelling process.

1. NYHA functional classification. Class I: patients with no limitation in ordinary physical activity, this includes asymptomatic left ventricular dysfunction. Class II: patients with slight limitation during physical activity and HF symptoms with regular physical activity. Class III: patients with marked limitation in physical activity, that is, asymptomatic at rest but with HF symptoms during any physical activity. Class IV: patients with HF symptoms even while at rest.
2. Score on the Minnesota Living with Heart Failure Questionnaire (MLHFQ), a Quality-of -Life Questionnaire specific to patients with HF. The MLHFQ is self administered, is made up of 21 items, yields a total score, and has two dimensions, physical and emotional. Items are scored from 0 to 5, and the result may be cited as a total score and/or as the scores of the two dimensions assessed, that is, physical (eight items) and emotional (five items). The Spanish version of the MLFHQ has been validated.[33]
3. WHODAS2.0 scale scores at the initial visit.

Initial visit

In the case of GP consultations, diagnoses will be confirmed and baseline situations evaluated, inclusion and exclusion criteria will be checked, Framingham criteria, NYHA functional classification and pharmacologic treatment will, likewise, be verified, and finally, an echocardiogram and proBNP determination will be requested, where necessary. Patients will be given appointments with a nurse for evaluation of quality of life (MLHFQ) and functional state (WHODAS 2.0).

Follow-up (visit 2)

Patients will be scheduled for evaluation by the GP and nurse, at 1 year of follow-up. In addition to a clinical examination similar to the initial visit (which does not include an echocardiogram), this evaluation will specifically check on disability and occurrence of hospitalisation or death.

All patients will be included in the health card database, thereby permitting them to be located even if they should change their GP.

Sample size

To predetermine the size of the sample, we used the possible relationship between experiencing a hospital admission in the first year of monitoring, as the dependent variable, and degree of disability, the explanatory variable, since more information was initially available on these two factors.

For calculation purposes, the following assumptions were made:

1. patients with a complete or severe degree of disability (exposed group) were assumed to have a relative risk of hospital admission in the first monitoring year of2.0when compared with patients with a moderate, mild or non-existent degree of disability (unexposed group);

2. the cumulative admission rate in the first year was assumed to be 15% in the unexposed group and, therefore, 30% in the exposed group;
3. accepting the population frequencies reported by de Pedro-Cuesta *et al*,[29] one would expect a population prevalence of complete or severe disability (exposed group) of 29.3% and, by extension, 70.7% in the unexposed group (null, mild or moderate disability). This would translate as an unexposed:exposed ratio of 2.41;
4. assuming an alpha risk of 0.05 and a beta risk of 0.2, then, in a two-tailed test, taking into account the fact that for each exposed individual, 2.41 unexposed individuals will be recruited, 109 individuals in the exposed group and 262 individuals in the unexposed group would be required in order to detect a minimum relative risk of two, if the proportion of patients in the unexposed group was 0.15. Losses to follow-up were estimated at 10%, using the GRANMO V.7 software program 12 (developed by URLEC Consortium of IMIM-Hospital del Mar, Barcelona, Spain);[34]
5. this sample size of 371 would be the minimum required to form the development sample for the predictive model. This size would allow for the development of logistic regression models with at least four explanatory variables;[35]
6. since one of the designated objectives is not only to develop but also to validate a predictive model, the size of the sample needed for validation purposes will have to be increased by 33% (186 individuals).[36] This sample will be completed once the development sample is fully formed;
7. to sum up, the total sample will comprise 557 patients. Considering the size of the target population and its usual prevalence rates,[1] a sample of this size could easily be recruited.

Plan of analysis

1. Quantitative variables will be described by calculating means, medians and dispersion of measures. Normality will be tested using the Shapiro-Wilk test: if the null hypothesis is rejected, values will be transformed to achieve adjustment. Qualitative variables will be described as proportions. CIs will be calculated for population inference purposes. Comparison of measures will be performed using the Student's t-test or analysis of variance: the association between qualitative variables will be studied using the χ^2 test or Fisher's exact test, and the association between quantitative variables will be studied using the Pearson correlation coefficient and its corresponding hypothesis test.
2. Construction of the predictive model: different models will be constructed using logistic regression. Predictive variables showing a p value of <0.25 in the crude analysis of association with the respective dependent variables will be used in the construction of the models, wherever this makes clinical sense. Thereafter, variables will be included having regard to their contribution to the model's significance, using the likelihood ratio test[37] and including the least number of variables so as to maximise usage in clinical practice. Possible interactions will be studied using the same criterion of inclusion as in the model. In view of their comprehensibility and potential for use in clinical practice, only possible interactions of order 2 which make clinical or biological sense will be studied. The models' goodness-of-fit will be evaluated using the Hosmer-Lemeshow test. The models will be constructed with a subsample (development sample) made up of a random selection of 70% of the total sample.
3. Model validation: the remaining individuals (30% of the sample) will be used for external validation of the different models. This will be performed using the model(s) selected to calculate the probability of the outcome of interest occurring for each individual included in the validation sample, when compared with the events actually observed to occur in this sample.[35] The χ^2 test will be used as the hypothesis test, since the aim of the models is to stratify patients according to their probability of presenting with the outcome of interest (hospitalisation during the course of the year).

As part of the validation process, the predictive capacity of each of the models will be estimated by calculating the receiver operating characteristic curves and the area under such curves, and if necessary, comparing the areas under the curves of each of the possible models. If the models prove valid and accurate enough for prediction purposes, scoring systems permitting a simpler use of the clinical practice models will be defined, using the method proposed by Spiegelhalter.[38]

All analyses will be performed using the STATA v.SE14 and R v.3.3 statistical software packages.

Study time frame

Patient recruitment began in July 2015 and is envisaged to be completed by March 2017. The follow-up period is scheduled to end in March 2018. The major part of the data analysis will be performed from April to July 2018.

DISCUSSION

HF is a chronic disease paradigm, involving a gradual increase in health resource requirement and the intervention of professionals from different healthcare and social levels. There is, therefore, a perceived need to develop a predictive model of undesirable events in primary care patients suffering from HF, which would allow for patients to be stratified according to their healthcare requirements. Such a model should ideally enable personalised attention, which would, in turn, lead to an improvement

in outcomes in terms of preventing hospitalisations and delaying disability.

The three fundamental elements of predictive model research in the prognosis research strategy are as follows: (1) model development, (2) external validation and (3) impact evaluation.[11] This study seeks to address the first two of these three elements.

Most of the published studies have sought to identify predictive factors of hospitalisation and death and have been based on data obtained from hospital admissions or hospital emergency departments. To date, demographic and clinical variables have mainly been considered as possible predictive factors, with largely inconsistent results. In our study, the fact that sample selection is to be based on primary care will ensure the inclusion of patients with a wide range of severity. We shall not only be trying to identify predictive factors of hospitalisation, both new admissions and readmissions, but also be seeking to identify predictive factors of functional impairment. In the models to be built in the future, we intend to include socioeconomic variables, quality-of-life variables and regarding the hospitalisation prediction model, variables of functional capacity, in addition to demographic and clinical variables.

The participation of GPs and nurses in the development of the proposed predictive model is likely to generate useful information for medical practice in primary care.

BIASES AND LIMITATIONS

For all intents and purposes, the study can be regarded as being population based, since it covers all segments other than patients who rely exclusively on private healthcare. The main data sources will be primary care EHRs and the population base having the right to public healthcare, which includes practically the whole population. Nevertheless, the population-based nature of the cohort is somewhat incidental, since the goal here is not to estimate population prevalence but to achieve a risk prediction. Classification bias will be minimised by performing echocardiograms on all patients who have not undergone a recent examination. While patients with HF in a very incipient stage may be left out of the study, it is improbable that such patients would develop the events of interest. Similarly, it is possible that some patients at an advanced stage of the disease might be excluded from the study because they suffer from an important functional limitation, but have not undergone an echocardiogram in the previous 6 months.

Due to its general nature, the socioeconomic variable used (average income in the census zone) may underestimate the contribution of individual socioeconomic characteristics.

Few withdrawals and losses to follow-up are expected, since the follow-up period is short. Any patients who withdraw their informed consent during the study will be asked for authorisation to conduct a search for their data in the electronic records.

ETHICS AND DISSEMINATION

This study will be conducted in compliance with the norms of good clinical practice and the principles laid down in the Helsinki Declaration. The project has been evaluated by the Madrid Regional Primary Care Research Committee and approved by the Clinical Research Ethics Committee of La Princesa University Teaching Hospital (PI-705). Informed consent will be requested from patients before their inclusion. Treatment of personal data will ensure that no information obtained can be associated with identified or identifiable persons, pursuant to the 1999 Personal Data Protection Act (*Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal*).

Results will be published in peer-reviewed journals and shared with medical community at conferences and scientific meetings.

Author affiliations
[1]Multiprofessional Education Unit for Family and Community Care (South-east), Madrid, Spain
[2]Research Network for Chronic Diseases (Red de Investigación en Servicios de Salud en Enfermedades Crónicas/REDISSEC), Madrid, Spain
[3]Subdirection for Infrastructures and Technology Assessment, Madrid Regional Health Authority, Madrid, Spain
[4]Cardiology Department, Henares University Teaching Hospital, Madrid, Spain
[5]Cardiology Department, La Princesa University Teaching Hospital, Madrid, Spain
[6]Telemedicine and e-Health Unit, Carlos III Institute of Health, Madrid, Spain
[7]Information Systems Department, Primary Care Management Division, Madrid, Spain
[8]Multiprofessional Education Unit for Family and Community Care (Centre), Madrid, Spain
[9]Carlos III Institute Agency for Health Technology Assessment (Agencia de Evaluación de Tecnologías Sanitarias-Instituto Carlos III/AETS-ISCIII), Madrid, Spain
[10]Ciudad San Pablo Health Center, Coslada, Madrid, Spain
[11]San Fernando Health Center, San Fernando de Henares, Madrid, Spain
[12]Los Alperchines Health Center, San Fernando de Henares, Madrid, Spain
[13]Jaime Vera Health Center, Coslada, Madrid, Spain

Correction notice This paper has been amended since it was published Online First. Owing to a scripting error, some of the publisher names in the references were replaced with 'BMJ Publishing Group'. This only affected the full text version, not the PDF. We have since corrected theseerrors and the correct publishers have been inserted into the references.

Collaborators CHIC Group: Cristina García Benito CGB, El Puerto Health Center, Coslada, Madrid; Cristina Gutiérrez Humanes CGH, El Puerto Health Center, Coslada, Madrid; José Sanz Arribas JSA, Los Alperchines Health Center, San Fernando de Henares, Madrid; Santiago Álvarez Arce SAA, Los Alperchines Health Center, San Fernando de Henares, Madrid; Marisol Blesa Mayordomo MBM, Doctor Tamames Health Center, Coslada, Madrid; Isabel Pedraz García IPG, Doctor Tamames Health Center, Coslada, Madrid; David García Gutiérrez DGG, El Puerto Health Center, Coslada, Madrid; Ángel Serrano Vázquez ASV, El Puerto Health Center, Coslada, Madrid; Nieves Gómez Escalonilla NGE, San Fernando Health Center, San Fernando de Henares, Madrid; María José Jarabo Sánchez MJJS, Doctor Tamames Health Center, Coslada, Madrid; Isabel Tello Meco ITM, Los Alperchines Health Center, San Fernando de Henares, Madrid; Esperanza Sánchez Rodríguez ESR, Jaime Vera Health Center, Coslada, Madrid; Pilar Aller Martín PAM, Los Alperchines Health Center, San Fernando de Henares, Madrid; Genoveva Delgado Pinto GDP, El Puerto Health Center, Coslada, Madrid; Amparo Martín Sanz AMS, El Puerto Health Center, Coslada, Madrid; Rosario Chantal Costero RCC, Doctor Tamames Health Center, Coslada, Madrid; Antonio Rodríguez Osorio ARO, Doctor Tamames Health Center, Coslada, Madrid; Eva Sevillano Fuentes ESF, Valleaguado Health Center, Coslada, Madrid; Belén Núñez Portela BNP, El Puerto Health Center, Coslada, Madrid; Mercedes Oliva Recio MOR, El Puerto Health Center, Coslada, Madrid; Remedios Ureña Llacer RUL, Doctor Tamames Health Center, Coslada, Madrid.

Contributors LG-O conceived the study and participated in its design and coordination. FR-S, MB-P, RA-T, CP-F, AG-C, MC, SR-A, SG-E, AA, LMS-G, RSM, EM-N, LB-O, NB-P, AS-P, ABLR, MAM-A, MAV-C, MIB-G and CGB participated in different phases of the protocol study design. LG-O wrote the final manuscript and FR-S, RA-T, MA, AA, SG-E, MC and LMS-G collaborated in the writing of the manuscript. All the authors and the authors from the CHIC group have read and approved the final manuscript.

Funding This study was funded by Health Research Fund (Fondo de InvestigacionesSanitarias/FIS) grant no. PI 14/01677 and co-financed with European Regional Development Fund (ERDF) funds (Carlos III Institute of Health-Research Network for Chronic Diseases/ISCIII–REDISSEC Project).

Competing interests None declared.

Patient consent Obtained.

Ethics approval Clinical Research Ethics Committee (CREC) of the University Hospital La Princesa.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Banegas JR, Rodríguez-Artalejo F, Guallar-Castillón P. Situación epidemiológica de la insuficiencia cardiaca en España. *Revista Española de Cardiologia* 2006;6(Suppl C):4–9.
2. Hospitalaria Edemorbilidad. Instituto Nacional de Estadística. *Madrid* 2011;2015 http://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176778&menu=ultiDatos&idp=1254735573175
3. Butler J, Kalogeropoulos A. Worsening heart failure hospitalization epidemic we do not know how to prevent and we do not know how to treat! *J Am Coll Cardiol* 2008;52:435–7.
4. Alonso J, Ferrer M, Gandek B, *et al*. Health-related quality of life associated with chronic conditions in eight countries: results from the International Quality of Life Assessment (IQOLA) Project. *Qual Life Res* 2004;13:283–98.
5. Naveiro-Rilo JC, Diez-Juárez MaD, Flores-Zurutuza L, *et al*. La Calidad de vida en los enfermos con insuficiencia cardiaca: visión desde atención primaria de salud. *Gac Sanit* 2012;26:436–43.
6. Mosterd A, Cost B, Hoes AW, *et al*. The prognosis of heart failure in the general population: the Rotterdam Study. *Eur Heart J* 2001;22:1318–27.
7. Carmona M, García-Olmos LM, Alberquilla A, *et al*. Heart failure in the family practice: a study of the prevalence and co-morbidity. *Fam Pract* 2011;28:128–33.
8. Carmona M, García-Olmos LM, García-Sagredo P, *et al*. Heart failure in primary care: co-morbidity and utilization of health care resources. *Fam Pract* 2013;30:520–4.
9. Lee WC, Chavez YE, Baker T, *et al*. Economic burden of heart failure: a summary of recent literature. *Heart Lung* 2004;33:362–71.
10. Poses RM, Smith WR, McClish DK, *et al*. Physicians' survival predictions for patients with acute congestive heart failure. *Arch Intern Med* 1997;157:1001–7.
11. Allaudeen N, Schnipper JL, Orav EJ, *et al*. Inability of providers to predict unplanned readmissions. *J Gen Intern Med* 2011;26:771–6.
12. Riley RD, Hayden JA, Steyerberg EW, *et al*. Prognosis research strategy (PROGRESS) 2: prognostic factor research. *PLoS Med* 2013;10:e1001380.
13. Dunlay SM, Redfield MM, Weston SA. Therneau TM, Long KH, Shah ND, Roger VL. hospitalization after heart failure diagnosis. A community perspective. *J Am Coll Cardiol* 2009;54:1695–702.
14. Kansagara D, Englander H, Salanitro A, *et al*. Theobald C, Freeman M, Kripalani S. Risk prediction models for hospital readmission: a systematic review. *JAMA* 2011;306:1688–98.
15. Ross JS, Mulvey GK, Stauffer B, *et al*. Statistical models and patient predictors of readmission for heart failure: a systematic review. *Arch Intern Med* 2008;168:1371–86.
16. Rahimi K, Bennett D, Conrad N, *et al*. Risk prediction in patients with heart failure: a systematic review and analysis. *JACC Heart Fail* 2014;2:440–6.
17. Hummel SL, Katrapati P, Gillespie BW, *et al*. Impact of prior admissions on 30-day readmissions in medicare heart failure inpatients. *Mayo Clin Proc* 2014;89:623–30.
18. Ather S, Chan W, Bozkurt B, *et al*. Impact of noncardiac comorbidities on morbidity and mortality in a predominantly male population with heart failure and preserved versus reduced ejection fraction. *J Am Coll Cardiol* 2012;59:998–1005.
19. Rodríguez-Artalejo F, Guallar-Castillón P, Pascual CR, Rodríguez Pascual C, *et al*. Health-related quality of life as a predictor of hospital readmission and death among patients with heart failure. *Arch Intern Med* 2005;165:1274–9.
20. Shih SL, Gerrard P, Goldstein R, *et al*. Functional status outperforms comorbidities in predicting acute care readmissions in medically complex patients. *J Gen Intern Med* 2015;30:1688–95.
21. Rathore SS, Masoudi FA, Wang Y, *et al*. Socioeconomic status, treatment, and outcomes among elderly patients hospitalized with heart failure: findings from the National Heart Failure Project. *Am Heart J* 2006;152:371–8.
22. Domínguez-Berjón F, Borrell C, Rodríguez-Sanz M, *et al*. The usefulness of area-based socioeconomic measures to monitor social inequalities in health in Southern Europe. *Eur J Public Health* 2006;16:54–61.
23. Lewis G, Curry N, Bardsley M. Choosing a predictive risk model: a guide for commissioners in England. *London. Nuffield Trust* 2011.
24. Murphy SM, Castro HK, Sylvia M. Predictive modeling in practice: improving the participant identification process for care management programs using condition-specific cut points. *Popul Health Manag* 2011;14:205–10.
25. Wong CY, Chaudhry SI, Desai MM, *et al*. Trends in comorbidity, disability, and polypharmacy in heart failure. *Am J Med* 2011;124:136–43.
26. Lupón J, González B, Santaeugenia S, *et al*. Implicación pronóstica de la fragilidad y los síntomas depresivos en Una población ambulatoria con insuficiencia cardiaca. *Revista Española de Cardiología* 2008;61:835–42.
27. Seo Y, Roberts BL, LaFramboise L, *et al*. Predictors of modifications in instrumental activities of daily living in persons with heart failure. *J Cardiovasc Nurs* 2011;26:89–98.
28. Vázquez Barquero JL, Herrera Castanedo S, Vázquez Bourgoón E, *et al*. Cuestionario para la evaluación de discapacidad de la Organización Mundial de la Salud (versión española del World Health Organization Disability Assessment schedule II). *Madrid. Ministerio de trabajo y asuntos sociales* 2006.
29. de Pedro-Cuesta J, Alberquilla Á, Virués-Ortega J, *et al*. ICF disability measured by WHO-DAS II in three community diagnostic groups in Madrid, Spain. *Gac Sanit* 2011;25(Suppl 2):21–8.
30. de Pedro-Cuesta J, García-Sagredo P, Alcalde-Cabero E, *et al*. Disability transitions after 30 months in three community-dwelling diagnostic groups in Spain. *PLoS One* 2013;8:e77482.
31. Ponikowski P, Voors AA, Anker SD, *et al*. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure. *Eur Heart J* 2016;2016:2129–200.
32. *The Johns Hopkins ACG System Version 11.0. technical reference guide*. Baltimore: The Johns Hopkins University, 2014.
33. Garin O, Soriano N, Ribera A, *et al*. Validación de la versión española del Minnesota living with Heart failure questionnaire. *Rev Esp Cardiol* 2008;61:251–9.
34. GRANMO. *Sample size and power calculator V 7.12*. Barcelona: Institut Municipal d'Investigació Médica. Available from. https://www.imim.cat/ofertadeserveis/sotware-public/granmo/ (accessed 13 October 2016).
35. Freeman DH. *Applied categorical data analysis*. New York, NY: CRC Press, 1987.
36. Altman DG, Vergouwe Y, Royston P, *et al*. Prognosis and prognostic research: validating a prognostic model. *BMJ* 2009;338:b605.
37. Hilbe JM. *Logistic regression models*. Boca Raton: Chapman & Hall / CRC, 2009.
38. Spiegelhalter DJ. Probabilistic prediction in patient management and clinical trials. *Stat Med* 1986;5:421–33.