A systems BIOlogy Study to TAilored Treatment in Chronic Heart Failure

Reporting

Project Information

BIOSTAT-CHF

Grant agreement ID: 242209

Project website

Project closed

Start date
1 April 2010

End date
31 March 2015

Funded under
Specific Programme "Cooperation": Health

Total cost
€ 15 618 635,20

EU contribution
€ 11 894 287,40

Coordinated by
ACADEMISCH ZIEKENHUIS GRONINGEN
Netherlands

Final Report Summary - BIOSTAT-CHF (A systems BIOlogy Study to TAilored Treatment in Chronic Heart Failure)

Executive Summary:
Heart failure is common in Europe and its prevalence is increasing as our population ages. Despite major improvements in care since the early 90s, it is still related to a poor prognosis. New therapies often failed, and obvious explanation being that the response to treatment is not homogenous. Treatment may need to be tailored to the individual patient. BIOSTAT-CHF aims to identify patients with a poor outcome, despite currently recommended therapy.

The major strength of BIOSTAT-CHF is its comprehensive approach to identify responders from non-responders, using clinical characteristics, genomics, transcriptomics, metabolomics, proteomics and biomarkers. These data were collected in 2509 patients from a large European study, conducted in 60 hospitals throughout 13 different countries in Europe (WP6). The second major strength of BIOSTAT-CHF that is collected the same information from a validation cohort of 1732 patients from Scotland (WP8).
WP2 was responsible for designing the clinical index study, and to provide detailed information about the studied population, endpoints, sample collections and a detailed dosing and uptitration medication scheme.
WP3 was responsible for the biomarkers collection and analyses. Novel candidate molecules were studied, such as Angiogenin, ANP, BNP, CRP, cystatin C, D-dimer, ESAM-1, galectin-3, GDF-15, LTBR, mesothelin, myeloperoxidase, neuropilin, NGAL, NT-proCNP, osteopontin, PCT, pentraxin-3, periostin, PIGR, MR-proADM, prosaposin B, RAGE, ST2, syndecan-1, TNFR-1, Troy, and VEGFR-1 have been measured.
In WP4, a Genome Wide Association Study (GWAS) was performed to identify genetic variances that contribute to the clinical response of patients included in BIOSTAT-CHF. From >800000 genetic variants, one variant on chromosome 11 showed an association at a genome-wide significant level (OR = 2.69 95%CI = (1.90 3.79) p = 1.86x10-8) with 12 further variants showing an association at P < 1x10-5). Identical GWAS analysis of the replication cohort yielded 13 variants with a P < 1x10-5.
In WP5, proteomics analyses were performed. Although single proteins were not strongly prognostic on its own (ROC AUC of individual MALDI peaks ranged from 0.51-0.55); the predictive value of a combination of proteins yielded much better predictive values (ROC AUC>0.8).
WP7 was designed to create and validate a risk prediction model for clinical outcome of patients with heart failure. The mean c-statistic value of all models combined was 0.66 (± 0.0005) with 0.71 (±0.001) , 0.63 (±0.001) and 0.68 for predicting mortality, hospitalization and mortality or hospitalization respectively.
Taken together, BIOSTAT-CHF has been successful in running two large-scale clinical trials in Europe, collecting extensive information, including GWAS, proteomics, metabolomics, transcriptomics, biomarkers and clinical phenotyping of more than 4000 patients with worsening heart failure, that had a very high clinical event rate. The first and main results already look very interesting, but it needs to be realized that the majority of information has yet to come out of this tremendous effort of a large group of researchers across Europe.

Project Context and Objectives:
Heart failure is common in Europe and its prevalence is increasing as our population ages. Despite major improvements in care since the early 90s, it is still related to a poor prognosis, an impaired quality of life and high health care costs. Many new therapies have failed to improve outcome further. One obvious reason is that the response to treatment is not homogenous. Treatment may need to be tailored to the individual patient. BIOSTAT-CHF aims to identify patients with a poor outcome, despite currently recommended therapy using information on demographics, gender, existing biomarkers, genetics and proteomics. Both genomic and proteomic analyses recently underwent major technical improvements, resulting in genome-wide analyses and detection of low abundance proteins. In BIOSTAT-CHF an index cohort of 2500 patients with signs of worsening heart failure will be recruited after initial stabilization. Treatment will be optimized according to the heart failure guidelines of the European Society of Cardiology with diuretics, ACE-inhibitors, betablockers and aldosterone antagonists. When patients are optimally treated, any change in symptoms and exercise tolerance will be evaluated. Patients will then be followed

up for a mean of 18 months, and mortality and heart failure hospitalizations will be recorded. By using a systems biology approach, incorporating information from demographic, biomarker, genomic, proteomic, and the initial response to therapy, a risk prediction model will be designed, identifying patients with a poor outcome on currently recommended therapy. This model will then be validated in a real-life cohort of 1800 heart failure patients with similar characteristics as the index cohort. BIOSTAT-CHF will therefore be a major step towards personalized medicine. Identifying patients with a poor outcome on currently recommended therapy might lead to further development of targeted therapies, eventually leading to improvements in outcome for patients with heart failure in Europe.

Index cohort
After approval by the EC, a detailed protocol and a synopsis have been made (WP2). The study was coordinated by the Study Coordinating Centre (SCC), who work with a so-called “hubs and spokes-model” (WP6). Approximately 48 centres from 8 different European countries have be approached to participate in the index-study. Every country leader was responsible for their own patient recruitment and quality of the data. The country leaders (hubs) were trained and checked by the SCC. The local investigators of the centres were the spokes. At baseline and after 6 months, blood samples (total of 70 cc) were drawn from the patients and stored in the freezer. These samples were transported to a central storage facility. The samples were then shipped in batches to the centres doing biomarker analyses (WP3), genomics (WP4), and proteomics (WP5). The outcome of biomarkers, genomics, and proteomics has been clustered with the patient characteristics (demographics)), and all data was electronically sent to the Systems Biology WP leader (WP7). This multilevel data was analyzed using logistic regression models, but also using alternatives, such as logic regression, neural networks, regression trees, random forests and support vector machines, to determine groups of patients that will or will not respond to chronic heart failure therapy.

Validation cohort
The risk model from the index cohort was then validated in a second cohort of 2500 chronic heart failure patients, who were recruited simultaneously. These patients were recruited in Scotland, UK, using extensive clinical NHS datasets in Tayside, administered by the Health Informatics Centre (HIC) at the University of Dundee, who provided an efficient research portal of anonymized data from all available clinical data resources in the region. The patients were similar to the index cohort, and the prediction model was adjusted to a final prediction model, to determine patients that do not respond to the current recommended therapy.
Project Results:
Work Package 2: Protocols
Its objective was to develop a detailed protocol for the index BIOSTAT-CHF study. This work package was subdivided into 10 tasks, shown below.
Task 1: to provide detailed information about study plan and procedures
A detailed protocol has been developed and distributed to the country Principal Investigators (PIs). The protocol was thoroughly discussed during a 2-days meeting on May 15-16th, 2010 with the participation of all the countries Principal Investigators (PIs). Further discussion followed by emails mainly regarding the inclusion and exclusion criteria and the characteristics of the study.
BIOSTAT-CHF is the acronym for “A systems BIOlogy Study to TAilored Treatment in Chronic Heart Failure”.

The rationale of this study was that, despite recent advances in treatment, the prognosis of patients with heart failure remains very poor. One likely cause of this is the use of homogeneous, standardized approaches exclusively based on the results of major clinical trials despite the heterogeneity in patients' clinical characteristics and responses to therapy.
On the other hand, due to major technological molecular oriented advances, first steps towards personalized medicine have recently been made through the implementation of biomarkers, genomics, or proteomics. BIOSTAT-CHF was aimed at the use of more updated methods for biomarkers, proteomic and genomic analyses and their integration with the demographic characteristics and available prognostic markers into a single model able to predict response to treatment and outcomes in each patient.
The aim of the study was to use a systems biology approach to evaluate which patient will have a poor clinical outcome despite evidence-based, heart failure treatment. Clinical, laboratory, genomic and proteomic data had to be developed and integrated into a risk score to identify patients with a poor response to treatment (death or heart failure hospitalization);).
Study design was that of a multicenter, multinational, prospective, observational study. Patients had to be recruited in approximately 10 different countries, in approximately 48 centers. The recruitment period was planned to be of 24 months, starting October 2010. Follow-up duration would have ranged from 18 months to 42 months
Study population had to include patients aged ≥18 years with symptoms of new-onset or worsening heart failure, confirmed either by a left ventricular ejection fraction of ≤40% or Brain Natriuretic Peptide (BNP) and/or NT-proBNP plasma levels >400 pg/ml or >2,000pg/ml, respectively, treated with either oral or intravenous (i.v.) furosemide ≥40 mg/day or equivalent at the time of inclusion, and who have not been previously treated with evidence based therapies (ACEi/ARBs and beta-blockers) or were receiving ≤50% of the target doses of these drugs at the time of inclusion and with an anticipated initiation or up-titration of ACEi/ARBs and/or beta-blocker therapy by the treating physician. Patients could be enrolled as in-patients or from out-patient clinics.
No specific intervention was forecasted. However, the treating physician was expected to optimize medical treatment according to the Heart Failure guidelines of the European Society of Cardiology(1) and good clinical practice recommendations.
Main study endpoints included as primary outcome death or unscheduled hospitalizations for heart failure. A patient with such an event was considered to be a non-responder and a patient without these events during the follow-up period was considered a responder. Number of events and days alive and out of hospital will be secondary outcome parameters.

Task 2: to provide a detailed dosing and uptitration medication scheme
the intervention in BIOSTAT-CHF consisted in the optimization of treatment of the patients who had developed worsening heart failure according to the current Heart Failure guidelines of the European Society of Cardiology(1) and good clinical practice recommendations by the treating physician. An intervention Dosing and uptitration scheme was already shown in the original protocol approved by the European Community (EC). As outlined in the original protocol, it was pointed out that an "anticipated initiation or up-titration of ACEi/ARBs and/or beta-blocker therapy by the treating physician" is an inclusion criterion.
Study design included an optimization phase, which would have directly followed the enrolment, during which "initiation or up-titration of ACEi/ARBs and/or beta-blockers will be done according to the routine clinical practice of the treating physician, who is expected to follow the current ESC guidelines."

Recommended target doses of ACEi/ARBs and beta-blockers were indicated in Table 1 and causes of lack of uptitration to target doses were shown in Table 2 of the protocol. Procedures for initiation and uptitration of therapy had to be the same as those outlined in the European Society of Cardiology (ESC) guidelines. (Dickstein K, Cohen-Solal A, Filippatos G, McMurray JJ, Ponikowski P, Poole-Wilson PA, Strömberg A, van Veldhuisen DJ, Atar D, Hoes AW, Keren A, Mebazaa A, Nieminen M, Priori SG, Swedberg K; ESC Committee for Practice Guidelines (CPG). ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2008: the Task Force for the diagnosis and treatment of acute and chronic heart failure 2008 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association of the ESC (HFA) and endorsed by the European Society of Intensive Care Medicine (ESICM). Eur J Heart Fail. 2008 Oct;10(10):933-89.)
The protocol was thoroughly revised in order to adhere to the characteristics of an observational and not of an interventional study. Thus, also with respect of drug dosing, it was pointed out that these procedures should have to follow the standards of good clinical practice with no specific intervention required.

Task 3: to provide a rationale for the studied population, endpoints, sample collections
Inclusion criteria.
Minor revision of the inclusion criteria were aimed at having a patient population the closest as possible to the characteristics of the general population of patients with heart failure while maintaining the initial characteristics of the study, as approved by the EC, i.e. the inclusion of patients with worsening symptoms of heart failure with an anticipated need of initiation or uptitration of ACEi/ARBs and/or beta-blockers. The main changes done to the protocol were the following:
• Patients with New onset or worsening heart failure (HF) can be enrolled. Differently from the initial protocol, an hospitalization for acute HF is not required as an inclusion criterion as also a patient undergoing an unplanned visit for worsening symptoms may be enrolled. This was also pointed out at the end of the inclusion criteria “Patient may be recruited as in-patients or from the out-patient clinic.”
• The study could include also patients with HF and preserved left ventricular ejection fraction (LVEF). With this aim, either a low LVEF or BNP and/or NT-proBNP plasma levels >400 pg/ml or >2,000 pg/ml, respectively, are used as inclusion criteria.

End-points
The primary outcome of interest was death or an unscheduled hospitalization for acute HF. A patient with such an event was considered to be a nonresponder and a patient without these events during the follow-up period was considered as a responder.
A detailed document with the definition of the end-points for adjudication was been prepared. It has been agreed that events will be adjudicated by each investigator under the responsibility of the country PI.

Samples collection
Blood samples for the assessment of patients’ biomarkers, proteomics and genomics had to be collected at baseline, at the time of entry into the study, and after 6 months of follow-up. Changes in these measurements would have allowed to assess the response to the changes in treatment done by the Investigator after the first episode of decompensation.
Standard Operating Procedures for sample collection and preparation for DNA and RNA analysis were defined and detailed in an ad-hoc document.

Task 4: to provide a detailed protocol for the 6 minutes walk test
A detailed protocol for the 6 minutes walking test was prepared, revised, and discussed. Similar to the documents above, this document was been circulated to the country PIs.

Task 5: to describe procedures and type of quality of Life questionnaire
The Kansas City Quality of Life Questionnaire and the EQ-5D score were adopted. They were shown in the appendix of the protocol and detailed manuals have been distributed. Questionnaires and their explanations awerere provided as appended files.

Task 6: to provide criteria and procedures for withdrawal of patients from the study
BIOSTAT-HF was an observational study. Thus, patient's withdrawal could only occur if she/ he withdraws her/ his consent to the participation to the study.

Task 7: to describe safety criteria and safety evaluation
Similar to above, as BIOSTAT-HF was an observational study a Data ad Safety Monitoring Board, as well as safety monitoring, were considered as not to applicable. As discussed in the protocol, safety reporting was organized at a single country level in case country's regulatory authorities consider safety reporting as needed also for an observational study.

Task 8: to provide definitions and reporting of (serious) adverse events
See above.

Task 9: to provide a generic patient information folder
A patient's consent form was prepared (see attachment)

Task 10: consensus and implementation
As above, a patient's informed consent form was prepared

Further developments
The original protocol was issued on 16 July 2010 and was amended on 22 February 2012. The amendment regarded mainly the possibility to recruit the patients as in-patients or from the out-patient clinic, the use of the most recent available LVEF determined within the prior 24 months and the use of BNP and/or NT-proBNP plasma levels measured during the current hospitalization or outpatient clinic visit. Further details which were amended regarded the blood sampling procedures.

Work Package 2: Biomarkers and Central Blood Bank
Its objective was to bundle logistics and the biomarker research related efforts in BIOSTAT-CHF and to develop a laboratory handbook for investigators.
A laboratory handbook with detailed instructions on sample collection, labelling, freezer system and shipment was prepared and was provided to each centre. In addition, all sites were trained for the procedures that related to the collection of the samples.
More than 5000 (serum and) EDTA plasma samples (2500 samples at two time points: baseline and 9 months visit) were collected from the project partners in Europe. We have started to study candidate molecules that may help in the characterization of heart failure patients with a poor outcome on currently

recommended therapy. B-type natriuretic peptide (BNP) is the standard biomarker in determining the diagnosis and prognosis of heart failure. Novel biomarkers, such as Angiogenin, ANP, BNP, CRP, cystatin C, D-dimer, ESAM-1, galectin-3, GDF-15, LTBR, mesothelin, myeloperoxidase, neuropilin, NGAL, NT-proCNP, osteopontin, PCT, pentraxin-3, periostin, PIGR, MR-proADM, prosaposin B, RAGE, ST2, syndecan-1, TNFR-1, Troy, and VEGFR-1 have been measured at baseline and after 9 months and analyses are currently under way.
Additional analyses are scheduled to be carried out including measurements of the following parameters at baseline and 9 months: troponin I, endothelin-1, BNP, TNF-alpha, interleukin-6, pro-enkephalin, Bio-ADM, aldosterone, plasma renin concentration, FGF-23, calcium, phosphate, and albumin.
The data is analysed with regards to their prognostic and diagnostic power, in particular with regards to the prediction of the clinical course of the patients, their hospitalisation rate, survival and with regards to the influence of co-morbidities.
The subgroup analysis will be performed in order to identify recently defined co-morbidities such as skeletal muscle wasting. Data will then be distributed to WP04 and WP10 partners.
We have started to study candidate molecules that may help in the characterization of heart failure patients with a poor outcome on currently recommended therapy. B-type natriuretic peptide (BNP) is the standard biomarker in determining the diagnosis and prognosis of heart failure. Novel biomarkers, such as Angiogenin, ANP, BNP, CRP, cystatin C, D-dimer, ESAM-1, galectin-3, GDF-15, LTBR, mesothelin, myeloperoxidase, neuropilin, NGAL, NT-proCNP, osteopontin, PCT, pentraxin-3, periostin, PIGR, MR-proADM, prosaposin B, RAGE, ST2, syndecan-1, TNFR-1, Troy, and VEGFR-1 have been measured at baseline and after 9 months and analyses are currently under way. Additional analyses are scheduled to be carried out including measurements of the following parameters at baseline and 9 months: troponin I, endothelin-1, BNP, TNF-alpha, interleukin-6, pro-enkephalin, Bio-ADM, aldosterone, plasma renin concentration, FGF-23, calcium, phosphate, and albumin.

Work Package 4: Genomics
Its objectives were to undertake a Genome Wide Association Study (GWAS) and to perform follow-on studies to identify genetic variances that contribute to the clinical response of patients included in BIOSTAT-CHF.

We have created and curated high-quality DNA (n = 2513) and a whole blood RNA (n = 2439) banks for the index BIOSTAT-CHF cohorts and a DNA bank from the replication cohort (n = 1654).
We have generated, quality-checked and curated genome-wide genotypes using the Affymetrix Axion UK Biobank array which includes probes for 825,928 SNPs for both the index and replication cohorts.
We have generated, quality-checked and curated genome-wide transcriptomic profiles for 945 selected subjects from the index cohort to facilitate downstream function analysis of associated variants and to identify expression profile signatures that may predict outcome.
We have undertaken genome-wide association (GWAS) analysis for the primary end-point of BIOSTAT-CHF (all-cause mortality and re-hospitalisation) in both the index and replication cohorts and identified putative novel genetic variants that may affect outcome. These are undergoing further validation.
We have provided the genomic data to WP7 for analyses using systems biology approaches.

Further details on the creation of the DNA and RNA banks are provided in Appendix 1 accompanying this report; on the performance, processing and quality control of the genotyping of the index cohort in

Appendix 2; on the GWAS analysis in the index cohort in Appendix 3 and on the genotyping and GWAS analysis in replication cohort in Appendix 4.
The DNA and RNA banks created in BIOSTAT-CHF are fully accessible to other investigators for additional analysis.

Work Package 5: Proteomics.
Its objectives were to undertake an examination of the proteomic profiles of plasma peptides/proteins that contribute to the variance in clinical response of patients to therapy in BIOSTATCHF, and to identify the principal peptide/protein components that will contribute to this variation in response.
The aim of the proteomics work package was to discover a panel of proteins or peptides that would help classify patients into those who responded well to drug therapy for heart failure (responders) and those who either died or were rehospitalised with heart failure (non-responders). Whilst many previous databases attempting to examine such problems were quite small, BIOSTAT-CHF provided a very large patient database in order to perform this task, so that all analyses were adequately powered.
Matrix assisted laser desorption and ionisation (MALDI) was used in this project, to generate ion profiles from extracts of plasma which could then be characterised on a high definition mass spectrometer. Many previous studies have utilised MALDI but usually in small sample sets, and also using instruments that were not high-resolution. Extraction of plasma on extra-wide pore reverse phase columns enabled us to study peptides up to a maximum molecular weight of about 8000, without the interference from high abundant proteins in plasma which normally mask the signals from low abundant proteins.
We obtained analysable spectra from 2248 patients with heart failure, of whom 1597 had no events within the first year (so called responders). The other patients (651) all had a hospitalisation with heart failure within that year, or had died. Using all the peptide intensity information in an individual patient's sample, it was possible to generate models that predicted death or death and/or hospitalisation that were at least 99% accurate. However, individual peptide peak intensities were not that informative, being accurate individually for predicting only 55% of events. This suggests that the whole spectra should be used whenever possible, as the summation of information from the whole spectrum was providing accuracy in prediction.
These results were internally cross-validated 10 fold, and yielded accuracies for death prediction of 88%, and for death and/or hospitalisation of 77%, suggesting that models may be applicable to other populations. The addition of clinical details of patients to this spectral information may further improve accuracies of prediction.
We were able to select over 100 spectral peaks that provided the most information for predicting the end points of death and death and/or hospitalisation. Due to the large number of peaks, it would not be practical to measure individual peptides in this spectrum to provide a measure of responder status in future. The method for generating a complete spectrum as a sort of bar code readout with intensity information may provide the basis for determining responder status in a high-throughput workflow. If this method was set up as a workflow for prediction of responder status, spectra could be added to the library in time, and endpoints obtained for any patient plasma studied, so that the growing database could become even more accurate. Our findings suggest high definition MALDI plasma profiling may have great potential to examine binomial questions in cardiovascular and other diseases, and may help classification of such patients without in-depth identification of the individual peaks.
The task of identifying single protein biomarkers from plasma, using specific antibody columns and peptide library beads to deplete plasma of high abundant proteins, led to a number of candidate proteins showing

differences according to responder status. However, on testing in larger patient cohorts, none of these candidates showed sufficient significant differences to assist in the classification of responder status.

Work Package 6: Patient related study
Its objectives were to create a state-of-the-art, collaborative, multinational, clinical research network of investigators who will work together to perform a clinical study develop an integrated score to identify patients more likely to show a favourable response to standard heart failure treatment. In order to do this, a large clinical study using the above mentioned infrastructure involving patients who are admitted for acute heart failure, were uptitrated to recommended heart failure therapy during 6 months, and were followed up for a median of 18 months to record death and heart failure hospitalisation.
Recruitment of the index study of BIOSTAT-CHF ended on 19 December 2012 and resulted in inclusion of 2519 patients. Patients follow-up has been completed in April 2014.
In total there were 1015 patients which had either a HF related admission or had deceased, which proved to be even higher than the anticipated event-rate, providing more than sufficient statistical power for the analyses.

Work Package 7: Systems Biology
Its objectives were to develop sparse matrix decomposition techniques to estimate the association of two or more high dimensional data sets. Second, the objective was to develop a predictive model for the response to heart failure therapy using baseline data; The third objective was to develop a prognostic model for the cardiac events occurring during the follow-up of the heart failure patients using baseline data and data about the response to therapy after six months; And the fourth objective was to validate this model in a validation cohort.
We conducted a systematic literature review to find which prediction models were already present in literature and which variables were used in those models.
We found 117 models developed to predict mortality, hospitalisation and/or mortality and hospitalisation. The prediction models were developed to predict outcome in a range of clinical settings. The model with the highest discriminatory power was used to predict in-hospital mortality, while the model with the lowest power predicted 6 months HF-hospitalisation. The mean c-statistic value of all models combined was 0.66 (± 0.0005) with 0.71 (±0.001) , 0.63 (±0.001) and 0.68 for predicting mortality, hospitalization and mortality or hospitalization respectively.
This meant that the models predicting mortality were more precise with respect to prediction than models predicting hospitalisation, with models predicting mortality and hospitalisation in between.
The models were developed in various different ways. Bigger models with more variables is had better predictive value. Predicting outcome early in the follow-up (1 year, e.g.) proved to be much easier than prediction late in the follow-up. Prediction models were most accurate when created with data from patients followed prospectively in a cohort study using data from medical records. Models predicting rehospitalization and mortality often used "administrative-claims-data" instead of data from medical records. These models predicting rehospitalization and mortality rates were however often developed for purposes different from those for individualized prediction of disease prognosis (like quantifying quality-of-care).
In the published models a large number of different variable were used. Most models used a combination of demographic, clinically and easily obtainable data to achieve the highest predictive power. The most frequently used variables, along with the number of times (#) used in the different models, and their

predictive power, are shown in table 1.

Predictive models

In the development of our prediction models for BIOSTAT-CHF we started with the 53 clinical, biochemical and demographic variables that were used in the existing prediction models. Missing values were imputed 5 times. We performed 1000 bootstrap backward selection analyses, separately for each of the outcome parameters. Our final model consisted of variables selected in more than 40% of the 1000 bootstrap samples. This procedure was repeated with the addition of 29 candidate biomarker variables measured by the Alere multiplex system and this resulted in 6 models. In two additional steps we included the first 10 principal components of the proteomic data and the first 10 principal components of the genetic data.
We evaluated these models based on their ability to correctly predict mortality using c-statistic values. In order to validate our model we conducted a bootstrap analysis where we corrected the c-statistic values for optimism. In this manuscript we present the internally validated optimism corrected data.
The first model with baseline clinical data and basic laboratory variables resulted in a model with 17 variables (NT-proBNP, Age, Urea, HB, LDL, Oedema Extent, HDL, Total Cholesterol, COPD, Sodium, DBP, SBP, LVEF, CABG, CAD, HT, PCI). This model had a c-statictic value of 0.74 (±0.005). With the Alere biomarker data added the c-statistic value improved to 0.749 (±0.005). This model existed of 15 variables (Age, Urea, NT-proBNP, ESAM1, ST2, HB, LDL, Total Cholesterol, WAP4C, PENTRAXIN3, HDL, SBP, Oedema Extent, COPD, CABG). Six variables (Sodium DBP LEVF CAD HT and PCI) were removed from the first model in advantage of ESAM1, ST2, WAP4C and PENTRAXIN3. The addition of proteomic and genetic data improved our mortality model to 0.751 (±0.004) and 0.756 (±0.004) respectively.
The baseline hospitalization existed of 9 variables (BNP, Age, HDL, NYHA class, DM, Urea, SBP, Race, LDL) and reached a c-statistic value of 0.694 (±0.004). The addition of biomarker data improved the c-statistics to 0.71 (±0.004). Here laboratory BNP was replaced by BNP measured by the Alere system and PENTRAXIN3 was the only other biomarker that was selected. The second model therefore consisted of 10 variables (BNP (Alere), PENTRAXIN3, Age, DM, Urea, HDL, NYHA class, Race, SBP, LDL). Addition of Proteomic Principal components made predictions worse. Genetic Principal components did improve c-statistic values compared to the third model, but did not improve c-statistic values compared to the second model.
The variables selected in the model predicting mortality and/or hospitalization existed of a combination of variables selected in the mortality and hospitalization models. 11 variables were selected in the baseline model (Age, Urea, NT-proBNP, HB, HDL, Sodium, NYHA class, Oedema Extent, SBP, CAD and COPD) with a c-statistic of 0.709 (±0.004). Adding alere data resuted in a model using 14 variables (Age, PENTRAXIN3, Urea, BNP (Alere), ST2, HDL, SBP, Oedema Extent, NYHA class, HB, CYSTATIN_C, COPD, Sodium, CAD). NT-pro-BNP was interchanged for BNP (Alere) and PENTRAXIN3, and CYSTATIN_C, were added, and no further variables were removed from the baseline model.

Validation of prognostic models

The prediction models were validated by ten-fold cross-validation. For every bootstrap-step, missing data were imputed five times. For every imputed set, the sample was divided at random into ten equal sized subsets. The prediction model was developed in nine of the ten subsets and validated in the tenth subset. This was repeated ten times such that every subset of the data was used as validation set. The c-statistics were averaged over subsets, imputation data sets and over the 1000 bootstrap samples. The cross-

validated c-statistics are reported in table 2. Compared to the averaged c-statistics of existing prediction models, the newly developed BIOSTAT-CHF prediction models have about 6% improved predictive value.

Work Package 8: Validation Cohort

Its objectives were to recruit an independent CHF population for validation and translation of the response-to-treatment model that will be generated from the index cohort. The second objective was to provide 'real world' automated longitudinal follow-up data through electronic record-linkage and tracking of all phenotypic endpoints [e.g. all drug use, all primary care consultations, ICD-9/10 coded hospital admissions including CHF hospitalisations, investigations, death (date and cause ICD 9/10)]. The third aim was to add value to total DNA and serum bioresource generated from this FP7 activity and to enhance management, governance and IT systems dedicated to this resource in order to maximize the international dissemination and utilization of the data and resource by all stakeholders

CHF patients were identified through a combination of criteria as originally defined in our study protocol:

- Previous hospital admission records for CHF based on ICD-9 code 428, ICD-10 code I-50, currently on combined CHF medications of loop diuretics and ACE inhibitors
- From echocardiogram database to identify patients with left ventricular dysfunction combined with drug profiling for CHF medications.
- Eligible patients who came through press release (Sept 2010).
- Eligible patients were identified from cardiology clinics

Our study protocol was submitted for ethics approval (R&D Ref Number 2008-CA03; MREC Number 10/S1402/39) on the 17th June 2010. Study protocol was approved on 10th July 2010.

As described in Task 8.2 we have successfully established originally planned recruitment centres to recruit patients identified to have CHF based on the case definition. These patients are recruited through the Tayside Health Informatics Centre with structured NHS shared care agreements that allow us to recruit patients either in hospital or in general practice care. Our original identified patient pool resides in Tayside and North-East Fife. Recruitment is by our assigned research nurse team at hospital sites in Tayside. In Tayside, the principal hospital sites are at Ninewells Hospital and Medical School, a major teaching hospital in Dundee and at Perth Royal Infirmary, a large district general hospital. At these sites, we have optimized screening strategies of all eligible CHF patients, including both inpatients and outpatients. This includes daily screening of all hospitalized CHF patients and of weekly cardiology/CHF clinics in Tayside. In the community, eligible patients identified to have CHF based on the predefined case definition are invited to participate in this study through the Tayside Health Informatics Centre with structured NHS shared care agreements.

All biological samples were collected, bar coded and delivered to the Biomedical Research Institute at Ninewells Hospital for sample processing and storage by a research technician and managed by our BIOSTAT laboratory manager (Dr Roger Tavendale). For patients recruited in sites outwith of Tayside, specimens were processed by a similar mechanism as established by close consultation with Dr Tavendale. All samples were logged into laboratory database using barcode scanners and the serum and plasma samples were spun and aliquoted into cluster racks, and stored immediately at -80c. Our freezers have dial out alarms and a dry ice protection system. Replicate aliquots were stored in separate freezers

for additional security. These samples are at this time still stored at the Biomedical Research Institute and will be sent to the respective WP partners for genomic, proteomic and biomarker analysis.
We have successfully provided access to the samples and dataset to our partners (WP3. CHARITE; WP4 ULEIC, WP5 ULEIC and WP7 AMC) in accordance to published access procedures. We have also established a link for both the distribution of samples/data to investigators and for the receipt of data from external and internal sources
At the end of recruitment period, we had recruited 1805 patients. Following subsequent data validation, there were 67 screen failure leaving a final cohort of 1738 patients available for analysis. The Clinical and Baseline Characteristics of this Final cohort (n=1738) are shown in table 3.

Through electronic record-linkage, we have successfully linked the patient datasets of the validation cohort to the Scottish national data held by electronic Data Research and Innovation Service (eDRIS) (http://www.isdscotland.org/Products-and-Services/eDRIS). eDRIS provides linkage to prescribing data, hospitalization data (including ICD-10 admission codes) from Scottish Morbidity Records (SMR01) and death data (including ICD-10 cause of death) from the General Records Office. This use of National Follow-up data enables continued follow-up of patients who move between regions within Scotland.

This was requested from eDRIS on 31/10/2014 therefore allowing us to capture at least a 6 month follow-up of the last patient recruited in April 2014. The application had to go through Privacy Advisory Committee (PAC) Approval to ensure protection of personal data and confidentiality and to ensure a safe haven for electronic health records (EHR) research. Data was received from eDRIS on 13/3/2015 and was loaded by Health Informatics Centre (HIC) based at the University of Dundee into their Safe Haven and we immediately assimilated this longitudinal follow-up data with the research datasets based in UMCG where the data is now held and made available for analysis

From the data available from eDRIS, we have a total follow-up of 3220 person years of follow-up time. 521 out of 1738 eligible cohort (30%) died up to Dec 2014. Death rate (95% CI) = 16.2 (14.8 17.6) deaths per 100 person-years Hospitalization data was complete till Oct 2104. 1337 people had 6,956 post-visit hospitalizations during follow-up. First Hospitalization rate (95% CI) = 103 (98-109) events per 100 person-years.

Potential Impact:
The overarching aim of BIOSTAT-CHF was to identify patients with heart failure that do not benefit from the currently recommended therapy. Patients were extensively characterized, including their full genetic make-up, their complete protein profile, including more than 50 biomarkers, and a large number of clinical characteristics. Taken together, BIOSTAT aimed to find a comprehensive and detailed picture of the “non-responsive patient”. The socio-economic impact is both direct and indirect. First, this patient population can become a target of upcoming drug development in heart failure. This would become a first step towards personalized/precision medicine, with obvious advantages. First, the outcome of the patients will improve. If we find better drugs that benefit patients and reduce the risk of hospital admissions, this will be good for the patients and for society. Second, unnecessary treatments will not be given to these patients, saving costs. Third, unnecessary side effects of the drugs can be avoided, improving the wellbeing of the patient, and further reducing costs. Therefore, the main outcome of BIOSTAT-CHF is a comprehensive risk prediction model, identifying patients that are treated according the current guidelines, but are still at

high to very high risk of dying or being hospitalized for acute decompensated heart failure. During the last 5 years, tremendous efforts have been made to provide the risk model. In the coming years, the results will be presented in large international congresses and international peer-reviewed publications in high ranked journals. Second, the careful determination of the profile of the non-responsive patient will have to lead to a better understanding of the disease, that will indirectly lead to better, and more targeted treatment options, further improving outcomes of patients with heart failure.

List of Websites:
www.biostat-chf.eu

Contact: Prof. Dr. A.A. Voors, University Medical Center Groningen, Dept of Cardiology, Hanzeplein 1, 9713 GZ Groningen, The Netherlands
Tel: +31 50 361 3238
Email: a.a.voors@umcg.nl

Related documents

final1-tables-final-report.pdf

Last update: 14 January 2016

Permalink: https://cordis.europa.eu/project/id/242209/reporting