A Remote Medical Monitoring System for Heart Failure Prognosis

Liangqing Zhang, 1 Cuirong Yu, 2 Chunrong Jin, 2 Dajin Liu, 2 Zongwen Xing, 3 Qian Li, 1 Zhinan Li, 4 Qin Li, 4 Yingxiao Wu, 5 and Jie Ren 6

1 Shanxi Cardiovascular Hospital, Taiyuan 030000, China
2 First Hospital of Shanxi Medical University, Taiyuan 030000, China
3 Taiyuan City Central Hospital, Taiyuan 030000, China
4 Taiyuan Maixinyun Healthcare Management Co. Ltd., Taiyuan 030000, China
5 Huaxin Consulting Co., Ltd., Hangzhou 310014, China
6 Shanxi Academy of Medical Sciences and Shanxi Dayi Hospital, Taiyuan 030000, China

Correspondences should be addressed to Yingxiao Wu; wuyingxiao@126.com and Jie Ren; renjie1011@163.com

Received 16 June 2015; Revised 25 August 2015; Accepted 14 September 2015

1. Introduction

Heart failure (HF), often known as congestive or chronic heart failure (CHF), is a common condition that develops after the heart becomes damaged or weakened by heart disease such as myocardial infarction, coronary artery, and rheumatic heart diseases [1]. HF is a life-threatening disease and addressing it should be considered a global health priority [2]. At present, approximately 26 million people worldwide are living with HF according to the statistics reported by Ponikowski et al. [3]. In many countries, population-based HF studies have shown that about 1 to 2% of people have HF and even higher proportions have been reported in single-centre studies [4]. Although many advanced techniques have been designed and used to treat HF, the risk of death is still about 35% in the year following diagnosis [5]. In China, according to the Chinese Cardiovascular Disease Report for 2013 [6], there were 4.5 million individuals living with HF and the morbidity was 0.9% for individuals between 35 and 74 years. The population of adults in the world with heart disease is continuously growing and ageing [7].

The signs and symptoms of HF are nonspecific (e.g., dyspnea, exercise intolerance, fatigue, and weakness) and often relate to other conditions such as pulmonary disease, anemia, hypothyroidism, depression, and obesity [8]. Several traditional tests and procedures are usually utilized for diagnostic assessment of HF such as blood test, B-type natriuretic peptide (BNP) [9] or N-terminal fragment of
the prohormone BNP (NT-proBNP) [10] test, electrocardiogram (ECG) [11], chest X-ray [12], echocardiogram [13], and coronary angiogram [14]. These methods require expertise from physicians, biologists, and clinicians. For example, NT-proBNP is an endogenously produced neurohormone primarily secreted by the ventricles in the heart as a response to left ventricular stretching or wall tension that occurs when HF develops and worsens. The level of NT-proBNP has been found to positively associate with age and negatively correlate to renal function [15]. No clear consensus has emerged for NT-proBNP as a diagnostic screening tool, but the age-adjusted cutting points (450 pg/mL for patients of <50 years, 900 pg/mL for patients of 50 to 75 years, and 1,800 pg/mL for patients of >75 years) appear promising and merit greater scrutiny and validation [16]. In clinical practice, the NT-proBNP test is considered an effective diagnostic method for HF and it provides guidance for HF therapy for patients with or without systolic dysfunction [17].

As known in the literature, HF is often a long-term (chronic) condition that generally worsens over time [18]. It requires frequent and costly hospitalization of patients for follow-up monitoring. As the number of patients with HF increases, clearly, there is a need to predict the occurrence of HF and thus provide early interventions with a home-based remote monitoring system to avoid long-term hospitalization or frequent NT-proBNP tests. Further, to reduce the morbidity and mortality of patients with HF and improve clinical outcome, its early prediction and prolonged monitoring are important in determining when to initiate specific therapies, in particular for patients with severe HF, such as cardiac transplantation and mechanical circulatory support [19]. Since therapies for HF are becoming more sophisticated, efforts are spent on developing more efficient and predictive models of HF risk and its changes so that personalized medical needs and goals of care for each patient are coordinated and communicated in order to provide the best solution of treatments. In addition, due to the inconvenience of geographic boundaries and/or economic constraints, monitoring heart health remotely appears to be a promising solution that can work at scale to improve HF prediction and reduce associated costs to both patients and hospitals. A major challenge to the realization of a remote monitoring system is the ability to collect, store, and process a large amount of data gathered from sensors in an effective, robust, and automated fashion. On the other hand, how to analyze the collected data to support the HF diagnosis or treatment is also a critical problem. The use of traditional NT-proBNP for timely diagnoses of HF usually requires a certain period of hospital stay (or frequent hospitalization) and it may not be appropriate for predicting impending HF during clinical follow-up.

To achieve remote monitoring of HF and potentially reduce patient hospitalization for testing, some physiological data that can be measured in a home environment using noninvasive sensors is required. For instance, Chaudhry et al. [20] showed that changes in body weight precede hospitalization for HF. Haider et al. [21] found that systolic blood pressure (SBP), diastolic blood pressure (DBP), and pulse pressure (PP, the difference between SBP and DBP) are predictors of risk for CHF. These studies point out that the use of body weight and blood pressure is promising for predicting future occurrence of HF and they are well-fitted to the requirements of our remote monitoring solution.

In this paper, we present the design and implementation of a remote medical monitoring system for HF prediction. The system design is an end-to-end solution including data collection, data storage and access, data analytics, and intervention feedback. The system provides prognoses of HF by estimating NT-proBNP level based on changes in blood pressure and body weight using machine learning methods, where a total of 29 features are extracted. To verify the effectiveness of the system, the NT-proBNP test is used as an aid to model the HF predictor estimates and to check the prognosis results. With this system, patients’ heart health can be remotely monitored.

The rest of this paper is organized as follows: Section 2 introduces related work on the remote monitoring for HF diagnosis/prognosis. Section 3 provides an overview of the system and its design principles. In Section 4, data collection, feature extraction and analysis, HF prediction algorithms, computation of a risk score for HF, and a pilot clinical trial are described. Section 5 presents the results of the prediction of HF and the computation of its risk score using our system and Section 6 discusses this study. Finally, Section 7 concludes this work.

2. Related Work

The conventional concept of a remote monitoring system in healthcare is often achieved by means of a telephone-based interactive voice response system. It usually collects patients’ daily information such as associated symptoms, feelings, and habits by asking questions. Afterwards, the collected information is then reviewed by clinicians. Chaudhry et al. [22] established a large trial on telemonitoring with 1653 patients enrolled, where 826 were randomly assigned to undergo telemonitoring and 827 to receive usual care. Telemonitored patients were required to make daily calls for six months. During each call, the patients were asked a series of questions about their general health and HF symptoms. The results showed that the telemonitoring strategy on a multicenter trial with a large database failed to provide a benefit over usual care and further strategies were needed to improve HF outcomes.

Suh et al. [23, 24] developed a system to remotely monitor patients with CHF, which had a three-tier architecture consisting of pervasive biosensors, a web server, and a back-end database. It acquired four health-related measures: weight, blood pressure, physical activity, and the Heart Failure Somatic Awareness Scale (HFSAS) [25] that reflect the most common signs and symptoms of CHF. Besides, their system can help the patients with guidance and feedback via text messages or emails.

Other researchers used a CardioMEMS [26] heart sensor in HF patients (Class-III according to the New York Heart Association [NYHA] guidelines [27]) to undertake a
single-blind trial [28]. The patients were managed with a wireless implantable hemodynamic monitoring system that collected pulmonary artery pressure. All patients in both treatment and control groups took daily pressure readings and then these measurements were transmitted through a modem to a database. Subsequently, the rate of HF-related hospitalizations was reduced by 37% in the treatment group. The potential limitations of the trial included the maintenance of patient masking and minimization of the effects of investigator-patient and device-patient interactions on HF outcomes.

A conceptual model for HF disease management (HFDM) was proposed in 2014 by Andrikopoulou et al. [29]. HFDM encompassed ongoing patient education and enhancement of self-care behaviour, complemented by data derived from device diagnostics. It had three basic mechanisms: (1) implementation of strategies that modify patients' baseline risk; (2) monitoring of worsening signs and symptoms; and (3) encouragement of patient participation in their own care. This model has been proposed as a way to systematically identify the risk factor for HF and, accordingly, prevent associated mortality.

Bui and Fonarow [30] reviewed some clinical trials and tested different HF monitoring strategies. They suggested several future challenges and opportunities in home-based hemodynamic monitoring such as evaluating HF monitoring in a broader population and in more diverse clinical settings, better defining optimal population for monitoring, studying long-term reliability and safety, and analyzing cost-effectiveness of home-based monitoring. Similarly, Bhimaraj [31] made an evidence-based review of various home-based monitoring systems for HF patients, including monitoring with telephone, portable technology, wearable sensors, and implantable cardioverter defibrillators (ICD) or cardiac resynchronization therapy (CRT) devices. They mentioned that the explosion of social media and smart-phone applications is a potentially untapped resource in creating a patient-centered system in the future.

Currently, many remote monitoring systems exist for HF assessment, where some of them use invasive and implanted sensors [28] and others monitor daily information to analyze HF via the Internet. Our proposed three-part monitoring system not only builds on existing technologies but also introduces new functions that circumvent their shortfalls. The remote system can achieve the diagnosis/prognosis of the occurrence of HF and provide interventions to patients for self-care treatment.

3. System Overview and Design Principles

3.1. System Architecture. The remote medical monitoring system for HF is built with three parts as shown in Figure 1. The first part is patient-end for acquiring data and sending/receiving feedback. It includes noninvasive sensing devices used for measuring body weight and systolic/diastolic blood pressure and a tablet with an “end-user” application (App) used for collecting questionnaire answers and interacting with patients. The second part is medical data center, to which data are sent from the patient-end. The medical data center stores the collected data, performs statistical data analysis and HF prediction, and generates data statistics and patient reports that can be delivered to doctors and patients. The third part is doctor-end, through which patient reports are sent to medical doctors who can then provide interventions/suggestions to patients according to their (current and previous) data statistics and reports.

These three parts form a circle so that the remote medical monitoring system can achieve early detection of HF as well as providing timely interventions. The system has several notable features. First, it can assess the patient’s condition by intelligently analyzing the physiological data. Second, it provides a real-time platform for asking and answering questions between doctors and patients. Third, it offers historical tracking for the patient’s heart health-related
data and medical records. Below are more details related to implementation of the system.

3.2. System Hardware and Software

3.2.1. System Hardware. As illustrated in Figure 1, the system includes three hardware devices:

(i) An electronic weight scale measuring body weight (kg).

(ii) An electronic sphygmomanometer measuring systolic blood pressure (SBP, mmHg) and diastolic blood pressure (DBP, mmHg).

(iii) A tablet “Heart-pad” which is a platform between patients and the system/doctors with an Android operation system including functions of collecting physiological data from the other two devices, collecting questionnaire data from patients, asking questions to doctors (by patients), and sending feedback to patients (by the system/doctors).

Note that all three hardware devices are connected using embedded Bluetooth 4.0 modules.

3.2.2. System Software. Several different types of software are implemented to achieve the prediction of HF occurrence. They are

(i) a date collection software in the tablet that serves to control the acquisition of body weight and blood pressure data and automatically upload the data to the medical data center;

(ii) a questionnaire generation software that automatically generates a certain number of questions for patients every day;

(iii) a prediction software that employs data mining algorithms (comprising feature extraction and structured support vector machine [SVM] classification) to analyze historical data and predict the future likelihood of the HF occurrence and thereafter computes a score indicating the HF risk (see Section 4.2);

(iv) a reporting software that visualizes some data and generates reports delivered to patients, where each report includes, for example, questionnaire statistics, daily changes of body weight and blood pressure (SBP and DBP), HF risk score (HFRS) variation, and suggestions by medical doctors and/or the system itself;

(v) a doctor-patient interaction software that provides a real-time platform for asking and answering questions between doctors and patients.

3.2.3. Graphical User Interface. In the system, we design two graphical user interfaces (GUIs), one for patients (patient-end) and the other for doctors (doctor-end).

(i) Patient GUI: Figure 2 illustrates some example pages of our patient GUI, implemented in an Android-based tablet. In the main page, patients can see different functions of the system and choose what they want to perform (Figure 2(a)). For example, they can measure their weight and blood pressure by following a voice instruction (Figure 2(b)), answer questions online (Figure 2(c)), or read the report...
with suggestions generated by the system or a doctor (Figure 2(d)).

(ii) Doctor GUI: some example pages are shown in Figure 3, where doctors can log in to the system through a desktop computer (Figure 3(a)), search and manage patients (Figure 3(b)), check the patients’ data (Figure 3(c)), and communicate with patients instantly where they may receive messages from their patients (Figure 3(d)).

4. Heart Failure Prediction and Heart Failure Risk Score

4.1. HF Prediction

4.1.1. Physiological Data. As mentioned above, we consider physiological data including body weight, SBP, and DBP on a daily basis for predicting the risk of occurring HF in consecutive 7. The prediction is based on modeling the objective collected during a 30-day period. To examine our prediction model, patients need to be classified as with or without a high risk of HF occurrence according to their age-adjusted NT-proBNP obtained on day 7 after 30 days of blood pressure and body weight monitoring. Therefore, the HF prediction is considered as a binary (predictive) classification problem, that is, classification of HF and non-HF patients using their past data. Figure 4 compares the body weight, SBP, and DBP for a period of 30 days for HF and non-HF patients. It illustrates that the variations of body weight, SBP, and DBP for HF patients seem greater than those for non-HF patients.

4.1.2. Questionnaire. In addition to the physiological data, our remote medical monitoring system also collects questionnaire data where 5 to 8 questions are sent to patients (i.e., to their Heart-pad terminal). The questions are selected from a large questionnaire pool of more than 100 questions. The questions are chosen by doctors or nurses according to the patient’s medical history. The patients are required to respond to them and the answers are then automatically uploaded to the medical data center. The use of a questionnaire is to assist doctors with understanding the objective data, provide suggestions, and make recommendations for further actions. Some example questions related to HF symptoms and identified as relevant based on medical data from clinical practice are shown as follows.

An Example of Questionnaire

1. Do you suffer from lack of appetite?
2. Do you sweat or feel nauseous?
3. Do you experience sudden confusion or have trouble speaking?
4. Do you experience severe headache?
5. Do you have shortness of breath (dyspnea)?
6. Do you feel fatigued?
7. Do you have physical restrictions or limitations?
8. Do you have symptoms of depression, loss of interest in activities, changes in sleep, loss of energy, and feelings of hopelessness?

4.1.3. Prediction Framework. With the acquired physiological data, we designed a framework to automatically predict/classify the occurrence of HF and test our classifier. It comprises a training and a classification stage. In the training stage, a set of features are extracted from training data and are used to train the classification model. In the
classification stage, features are extracted and classification is performed based on the trained model. After that, results are compared with the annotations of the test samples. The prediction framework with different blocks is illustrated in Figure 5.

4.1.4. Features Extraction and Discriminative Capability. In order to predict the occurrence of HF, we extract a total of 29 features that are characteristic for HF. They are extracted from body weight, SBP, DBP, and PP (pulse pressure, computed as the difference between SBP and DBP) values over a certain period. Table 1 lists all of the 29 features and definitions.

It is important to understand which features are more informative in detecting HF. To quantify the feature discriminative capability, a Mahalanobis distance (MD) metric [32] can be used. For a single feature, it computes the absolute standardized mean difference between two classes. The formula for calculating MD is given by

\[
MD = \frac{|\mu_p + \mu_n|}{\sigma},
\]  

Figure 4: Comparison of body weight, SBP, and DBP measures of 30 days for non-HF and HF patients.

Figure 5: HF prediction framework.
Table 1: List of features for HF prediction.

<table>
<thead>
<tr>
<th>Index</th>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$W_{\text{mean}}$</td>
<td>Mean weight</td>
</tr>
<tr>
<td>2</td>
<td>$W_{\text{max}}$</td>
<td>Maximum weight</td>
</tr>
<tr>
<td>3</td>
<td>$W_{\text{std}}$</td>
<td>Std. of weight</td>
</tr>
<tr>
<td>4</td>
<td>$W_{\text{inc}}$</td>
<td>Daily weight increase $&gt; 1$ kg</td>
</tr>
<tr>
<td>5</td>
<td>$W_{\text{incmax}}$</td>
<td>Daily weight increase $&gt; 1.5$ kg</td>
</tr>
<tr>
<td>6</td>
<td>$W_{\text{in}}$</td>
<td>Mean of daily weight increase</td>
</tr>
<tr>
<td>7</td>
<td>$W_{\text{instd}}$</td>
<td>Std. of daily weight increase</td>
</tr>
<tr>
<td>8</td>
<td>$W_{\text{inmax}}$</td>
<td>Maximum of daily weight increase</td>
</tr>
</tbody>
</table>

Systolic blood pressure (SBP) features

| 9     | $\text{SBP}_{\text{mean}}$ | Mean SBP |
| 10    | $\text{SBP}_{\text{max}}$ | Maximum SBP |
| 11    | $\text{SBP}_{\text{min}}$ | Minimum SBP |
| 12    | $\text{SBP}_{\text{std}}$ | Std. of SBP |
| 13    | $\text{SBP}_{\text{max}}$ | SBP $> 200$ mmHg |
| 14    | $\text{SBP}_{\text{min}}$ | SBP $> 160$ mmHg |
| 15    | $\text{SBP}_{\text{min}}$ | SBP $< 100$ mmHg |
| 16    | $\text{SBP}_{\text{max}}$ | SBP $< 90$ mmHg |
| 17    | $\text{SBP}_{\max}$ | SBP $< 85$ mmHg |
| 18    | $\text{SBP}_{\text{inc}}$ | Mean daily SBP increase |
| 19    | $\text{SBP}_{\text{instd}}$ | Std. of daily SBP increase |

Diastolic blood pressure (DBP) features

| 20    | $\text{DBP}_{\text{mean}}$ | Mean DBP |
| 21    | $\text{DBP}_{\text{max}}$ | Maximum DBP |
| 22    | $\text{DBP}_{\text{min}}$ | Minimum DBP |
| 23    | $\text{DBP}_{\text{std}}$ | Std. of DBP |
| 24    | $\text{DBP}_{\text{inc}}$ | Mean daily DBP increase |
| 25    | $\text{DBP}_{\text{instd}}$ | Std. of daily DBP increase |

Pulse pressure (PP) features

| 26    | $\text{PP}_{\text{mean}}$ | Mean PP |
| 27    | $\text{PP}_{\text{std}}$ | Std. of PP |
| 28    | $\text{PP}_{\text{corr}}$ | Corr. between SBP and DBP |
| 29    | $\text{PP}_{\text{incorr}}$ | Corr. between daily SBP and DBP increase |

where $\mu_p$ is the mean of positive class, $\mu_n$ is the mean of negative class, and $\sigma$ is the population standard deviation. A larger MD means that the feature is more informative in separating the two classes. Note that we consider HF and non-HF as positive and negative class, respectively.

4.1.5. Prediction Model. As stated, a well-known structured SVM classifier [33] is employed to classify HF and non-HF patients in this work. The training principle behind SVM is to find the optimal linear hyperplane such that the expected classification error for unseen samples is minimized. Here the SVM kernel for classification is a polynomial with order 3. More details about the structured SVM classification algorithms can be found elsewhere [33].

4.1.6. Evaluation Metrics and Cross-Validation. The evaluation metrics of classification performance used in this work include overall accuracy, sensitivity (or recall), specificity, and precision (or positive predictive value, PPV). Overall accuracy is computed as the ratio of correctly classified patients to the total number of patients; specificity is a measure that indicates the proportion of correctly classified actual negatives; sensitivity is the proportion of correctly classified actual positives; and precision is computed as the ratio of true positives to true positives plus false positives. These metrics can be computed from a confusion matrix (Table 2) generated based on the classification results such as

(i) accuracy = ($TP + TN$)/($TP + TN + FP + FN$) × 100%;
(ii) sensitivity = $TP$/$TP + FN$ × 100%;
(iii) specificity = $TN$/$TN + FP$ × 100%;
(iv) precision = $TP$/$TP + FP$ × 100%.

Additionally, a receiver operating characteristic (ROC) curve (plotted as sensitivity versus one minus specificity) of the results can be used to offer an overview of the classifier’s performance. In general, a larger area under the ROC curve (AUROC) corresponds to a better performance.

To validate the classifier without biasing the prediction results, we apply a leave-one-out cross-validation (LOOCV) with 34 iterations. During each iteration of the LOOCV procedure, 33 patients are used to train the classifier and the remaining one is used for testing. The above-mentioned evaluation metrics are then computed based on the classification results after LOOCV.

4.2. Heart Failure Risk Score. For the purpose of providing meaningful information to patients regarding their HF risk, we create a scoring system that is able to quantify the risk of the future occurrence of HF, yielding an HF risk score (HFRS). In fact, the structured SVM classifier also delivers posterior probability when assigning classes for decision-making. This posterior probability indicates the probability of a sample being classified as a specific class, which associates with the likelihood of the occurrence of HF. Therefore, this output can then be used to generate a score that indicates the HF risk.

This work proposes to determine HFRS values based on the analysis of classification results (i.e., specificity and sensitivity). We use a simple piecewise linear function with thresholding three “levels” (1 to 3) to map a posterior probability into an HFRS ranging from 0 to 100. The posterior probability range of the first level should ensure that all the patients with their posterior probabilities falling in this range are correctly classified as non-HF, indicating a low risk of HF. Similarly, the posterior probability range of the third
level should make sure that all the patients who have their posterior probabilities in this range are correctly classified as HF, indicating a high risk of HF. The other posterior probability values outside the first and the third level ranges are in the second level, indicating a moderate risk of HF where misclassifications may occur. The purpose of using an HFRS range between 0 and 100 is to provide a comprehensive score to patients so that they can easily understand and assess their risk level.

4.3. A Pilot Study. To evaluate our system, a pilot clinical trial was performed based on a total of 34 Chinese patients (18 females, age 67.2 ± 8.3 years) who have been clinically diagnosed to have heart disease (coronary artery, pulmonary, or rheumatic heart disease, dilated cardiomyopathy, cardiac arrhythmia, or myocardial infarction). The confirmation of heart disease diagnosis was done through blood tests, chest X-ray, echocardiogram, ECG, ejection fraction, angiogram, cardiac computerized tomography scan, and/or magnetic resonance imaging. In clinical practice, the patients were classified as having different severity levels of heart function (from stage-A to stage-D) according to the updated guidelines of the American College of Cardiology/American Heart Association (ACC/AHA guidelines) [34], where one patient was in stage-A (at high risk for HF in the future but no structural heart disease or HF symptoms), seven were in stage-B (with structural heart disease but no HF signs or symptoms), 19 were in stage-C (underlying structural heart disease with previous or current HF symptoms), and seven were in stage-D (refractory HF requiring specialized inventions or hospital-based support).

In this pilot study, we continuously measured the patients’ body weight, SBP, and DBP as well as questionnaire data on a daily basis for 30 days, where the measurements were performed before their breakfast in the morning in order to minimize the effect of food intake. Patients were then asked to measure their NT-proBNP value in a hospital 6 days later (on day 7), yielding annotated 22 HF and 12 non-HF patients. The data was collected in Shaxi Cardiovascular Hospital, Taiyuan, China, during the period between April and August, 2014. The devices in our remote medical monitoring system with installed software were provided by Sennotech Inc., China [35].

Here we considered two prediction schemes including predicting HF based on the physiological data during the past 30 days and during the past 7 days. This served to investigate the effect of using different historical data lengths on the ultimate classification performance. In other words, we intended to know if we could obtain acceptable prediction results when patients participated their remote monitoring for only a week. As mentioned, the prediction results were verified by an NT-proBNP test. The purpose of this pilot study was to preliminarily examine the usefulness of our system from HF prediction and implementation perspectives.

<table>
<thead>
<tr>
<th>Data</th>
<th>Precision</th>
<th>Sensitivity</th>
<th>Specificity</th>
<th>Overall accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>7 days</td>
<td>83.4%</td>
<td>63.6%</td>
<td>75.0%</td>
<td>67.6%</td>
</tr>
<tr>
<td>30 days</td>
<td>83.3%</td>
<td>77.3%</td>
<td>89.5%</td>
<td>79.4%</td>
</tr>
</tbody>
</table>

Figure 6 shows the discriminative capability (as measured by Mahalanobis distance [MD]) of all the 29 physiological features (computed based on past 30 days). It indicates that the mean and maximum body weights, the daily weight increases larger than 1 and 1.5kg, the number of SBP larger than 100mmHg, and the mean and standard deviation of DBP have higher discriminative capabilities than the other features. This is because patients with a higher weight, a larger daily weight increase, and larger variations in blood pressure will have a higher possibility for HF to occur after a week.

As shown in Table 3, our system with an SVM classifier achieved an HF prediction accuracy of 79.4% when using the daily weight and blood pressure measures collected during the past 30 days, where the specificity, sensitivity, and precision are also presented. However, the accuracy decreased to 67.6% when only using the past 7 days’ data, which indicates that including more historical data with a longer period can improve the prediction performance. This can also be observed by looking at their ROC curves (Figure 7), where the curve obtained with 30 days’ data has a larger “area under the ROC curve” compared with using data from the past 7 days.

As stated before, the HFRS scoring can be implemented by converting the SVM posterior probability outputs to HFRS values via a piecewise linear mapping method. Based on the posterior probabilities and the prediction results, we generated the linear mapping for the three different levels in the following. For the first level (Level 1), the posterior probability from 0 to 0.15 was linearly mapped to HFRS from 0 to 50, where the specificity was 100%. Patients classified at this level have a low risk of HF and they are
Table 4: Heart failure risk score.

<table>
<thead>
<tr>
<th>Level</th>
<th>Posterior probability</th>
<th>HFRS</th>
<th>Status of weight</th>
<th>Status of blood pressure</th>
<th>Suggestions or interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 low risk</td>
<td>0–0.15</td>
<td>0–50</td>
<td>Small variability in weight change and daily weight increase/loss</td>
<td>Small variability in SBP and DBP changes and their daily increases; normal PP (around 40 mmHg)</td>
<td>Keep on monitoring and provide lifestyle-related suggestions (e.g., improving living habits) according to questionnaire data</td>
</tr>
<tr>
<td>2 moderate risk</td>
<td>0.15–0.25</td>
<td>50–80</td>
<td>Moderate and increasing variability in weight changes and daily weight increases</td>
<td>Moderate and increasing variability in SBP and DBP changes and their daily increases; occurrence of some large daily increases in SBP or DBP; relatively low PP (&lt;40 mmHg)</td>
<td>Suggestions and interventions are needed from doctor after checking patient's physiological statistics and questionnaire data</td>
</tr>
<tr>
<td>3 high risk</td>
<td>0.25–1</td>
<td>80–100</td>
<td>Large variability in weight change and daily weight increase/loss</td>
<td>Large variability in SBP and DBP changes and their daily increases; occurrence of many large daily increases in SBP or DBP; fast SBP drops; low PP (&lt;20 mmHg) and large PP variability</td>
<td>Strongly suggest patients to be hospitalized</td>
</tr>
</tbody>
</table>

Figure 7: ROC curves of HF prediction using 30 days’ and 7 days’ data.

In our clinical trial, we executed a pilot study with only 34 patients. This small data set might lead to limitations on getting an accurate HF predictor and adequate ranges for computing HFRS levels. Therefore, an expanded clinical trial with more patients involved is necessary for future studies. Additionally, including a longer period of data (e.g., 3 to 6 months) may help improve the prediction results and the number of days is needed to obtain a converged performance.

Although the prediction results (Table 3) are far better than random guess, they are still under our expectations. Extracting more advanced features that can express more HF-related pathophysiological information is ideal in predicting its future risk, such as nonlinear entropy-based measures for single source [36] and multiple sources [37]. Here the SVM-based algorithm was used to perform HF and non-HF classification. Although it is an advanced algorithm and has been applied successfully in many different areas, other classifiers (e.g., logistic regression, random forest, and neural networks) still merit further investigation and comparison.

Our monitoring system requires patients to interact with it regularly (daily). Although the data collection and transmission are automatic with patient input, it will still be a challenge for patients to comply even if they have risk of HF. Patient compliance with our system should be further studied and analyzed. To reduce variations from patient to patient and between days when collecting their physiological data, we required them to measure their data at a fixed time (i.e., in the morning before lunch). However, this may result
HFRS report

Name: Shengzhou Guo  Gender: Male  Age: 72  ID: 140102194005271819

Did you drink last week? Yes: 1 time, no: 6 times

How is your aerobic exercise frequency (e.g., walking, running, bicycle riding) last week? 2–4 times a week

Did you feel dizzy last week? Yes: 3 times, no: 4 times

Did you feel short of breath last week? Yes: 2 times, no: 5 times

HRFS Index variation

Dynamic data

Doctor suggestion

(I) Reduce the use of high-fat food and sweet food.
(II) Quit smoking and wine.
(III) Keep routine and a good mood.

No.: XXXX1404140001

FIGURE 8: An example of HFRS report (translated from Chinese to English).
in inconvenience for patients or patient failure to measure data, therefore yielding missing data. Offering alternatives for the time of day patients which are required to collect data may help resolve this time effect in our prediction model. For this purpose, using a larger data set with more data for each patient measured at different times is suggested for future studies.

As stated, this present study only focused on HF for patients with diagnosed heart diseases. However, with our system, it is still unclear how to simultaneously process all the clinical data, many of which might lead to potentially conflicting alerts when the alerts relate to various unexpected comorbid conditions, resulting from medications. For example, for elderly patients with atrial fibrillation and concomitant HF, a rate-limiting calcium channel blocker would be effective for the first condition but potentially dangerous for the second one.

The future occurrence of HF in this study was estimated with NT-proBNP text. This would not be the most accurate parameter as a “ground-truth” of the occurrence of HF. Future study must validate our system based on the historical noninvasively measured data (body weight and blood pressure) for patients who actually have HF rather than using the NT-proBNP-based parameter that only indicates HF risk.

Finally, since the HFRS is based on the HF prediction results which are a data-driven scoring method, it needs to be further validated with more long-term clinical data and corrected by doctor’s expertise in clinical practice. As mentioned, the patients enrolled in this study included those with several different types of heart diseases which should be taken into account during HF prediction in future work. In addition to the heart disease type, the patients might not statistically represent the whole population, where ethnic group and patient demographics (e.g., age and gender) would likely influence the prediction models. In order to improve the HF prediction performance and to provide patient-specific interventions by achieving a personalized (prediction and intervention) system, analyzing a broader range of more ethnically diverse groups needs to be further investigated.

7. Conclusion

An effective remote medical monitoring system for heart failure (HF) prediction and management was designed and implemented. The system realized early prediction (or prognosis) of future HF occurrence by estimating future NT-proBNP level based on a patient’s historical data (body weight and blood pressure), where data were obtained remotely using noninvasive devices (i.e., a Bluetooth-based weight scale and sphygmomanometer). A Heart Failure Risk Score (HFRS) was proposed to evaluate the risk of the future occurrence of HF based on the prediction results, where the HFRS scoring was designed so that it could be easily understood and perceived by patients. This system optimizes early stage delivery of multiple suggestions/interventions to patients in different risk levels. This end-to-end system can also be used to manage patients and their data by doctors or by the system’s data center. To validate our system, a set of real-life data from 34 patients was collected in a pilot clinical trial. Our HF prediction algorithms achieved an overall accuracy of 79.4% and 67.6% when using the data collected over a 30-day and 7-day period, respectively. Therefore, from a clinical perspective, this system is promising to help reduce morbidity and mortality caused by HF and therefore improve the clinical outcome.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

References


