Original Paper

Early Indication of Decompensated Heart Failure in Patients on Home-Telemonitoring: A Comparison of Prediction Algorithms Based on Daily Weight and Noninvasive Transthoracic Bio-impedance

Illapha Cuba Gyllensten^{1,2}, MSc; Alberto G Bonomi¹, PhD; Kevin M Goode³, PhD; Harald Reiter¹, MSc; Joerg Habetha¹, PhD; Oliver Amft^{2,4}, PhD; John GF Cleland⁵, MD, PhD

¹Personal Health Solutions, Philips Research, Eindhoven, Netherlands

²Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, Netherlands

³Department of Health Professional Studies, Faculty of Health & Social Care, University of Hull, Kingston-Upon-Hull, United Kingdom

⁴ACTLab, University of Passau, Passau, Germany

⁵National Heart & Lung Institute, Imperial College, London, United Kingdom

Corresponding Author:

Illapha Cuba Gyllensten, MSc Personal Health Solutions Philips Research 5.007 High Tech Campus 34 Eindhoven, 5656AE Netherlands Phone: 31 631926930 Fax: 31 40274276 Email: <u>illapha@gmail.com</u>

Abstract

Background: Heart Failure (HF) is a common reason for hospitalization. Admissions might be prevented by early detection of and intervention for decompensation. Conventionally, changes in weight, a possible measure of fluid accumulation, have been used to detect deterioration. Transthoracic impedance may be a more sensitive and accurate measure of fluid accumulation.

Objective: In this study, we review previously proposed predictive algorithms using body weight and noninvasive transthoracic bio-impedance (NITTI) to predict HF decompensations.

Methods: We monitored 91 patients with chronic HF for an average of 10 months using a weight scale and a wearable bio-impedance vest. Three algorithms were tested using either simple rule-of-thumb differences (RoT), moving averages (MACD), or cumulative sums (CUSUM).

Results: Algorithms using NITTI in the 2 weeks preceding decompensation predicted events (P<.001); however, using weight alone did not. Cross-validation showed that NITTI improved sensitivity of all algorithms tested and that trend algorithms provided the best performance for either measurement (Weight-MACD: 33%, NITTI-CUSUM: 60%) in contrast to the simpler rules-of-thumb (Weight-RoT: 20%, NITTI-RoT: 33%) as proposed in HF guidelines.

Conclusions: NITTI measurements decrease before decompensations, and combined with trend algorithms, improve the detection of HF decompensation over current guideline rules; however, many alerts are not associated with clinically overt decompensation.

(JMIR Med Inform 2016;4(1):e3) doi: 10.2196/medinform.4842

KEYWORDS

RenderX

Heart failure; telemonitoring; deterioration detection; alert algorithms; ambulatory monitoring; impedance

Introduction

Chronic heart failure (HF) is common [1] and a substantial drain on scarce healthcare resources [2]. Much of the costs of HF are due to the high rate of unplanned admissions for worsening HF. For patients who survive an admission for worsening HF, rehospitalization rates are high and >20% will die within one year [3,4]. Furthermore, the high prevalence and costs associated with HF are projected to rise as the population ages [5]. Telemonitoring could reduce costs and improve outcomes [6] by substituting infrequent assessments at a clinical facility by a health professional with frequent remote monitoring done by patients themselves. This could facilitate more timely and tailored interventions. The efficacy of telemonitoring would be greatly improved if decompensation events could be detected before the onset of severe symptoms [7,8].

Worsening heart failure may lead to weight gain as a consequence of fluid retention and edema and, if uncorrected, can lead to hospitalization and ultimately death. The Heart Failure Association of America (HFSA) and the European Society of Cardiology (ESC) guidelines both recommend daily weight monitoring. The ESC recommends that patients experiencing a weight increase of 2 kg or more in 3 days should alert healthcare professionals and increase their diuretic dose [9]. The HFSA recommends the restriction of sodium and water after an increase of more than 2 lbs (0.9 kg) in 1 day, or more than 4 lbs (1.8 kg) over a week, followed by an alert to healthcare professionals if the increase continues [10].

Worsening hemodynamics with increased vascular resistance, afterload mismatch, congestion, and diastolic dysfunction are thought to precede fluid accumulation [11]. Increased end-diastolic pulmonary arterial pressure (PAP), a direct measure of hemodynamic overload, and decreased intrathoracic impedance (ITI), an indirect measure of pulmonary congestion, have both been observed in the days and weeks prior to decompensation [12-14]. Thoracic impedance can also be measured noninvasively (NITTI) [15], which correlates with ITI [16], making measurement possible in a far broader range of patients. NITTI measures a much larger field; however, the variability in measurements may depend on the patients' willingness and ability to position electrodes accurately. Recently, several new wearable devices have been proposed for this purpose, such as specialized vests [17,18] or adhesive patches [19,20].

An increased risk of decompensation has been shown for both weight gain [21] and decline in ITI [22]; however, recent studies have shown that absolute changes in weight over short time periods are not sensitive in detecting impending decompensation [23-25], and that ITI may have high sensitivity but a high rate of false alarms per patient-year [26]. However, to the authors' knowledge, recently proposed prediction algorithms comparing body weight and impedance head-to-head have not been investigated using *noninvasive* technology.

The aim of this investigation was to evaluate and compare the predictive value of previously published algorithms using measurements of daily body weight, and noninvasive measures of NITTI from a smart-textile vest, to detect decompensation prior to the onset of severe symptoms leading to hospitalization.

Methods

Patient Population

The data for this analysis were collected as part of the MyHeart heart failure management observational study [27]. The MyHeart study was unique in its collection of several different vital signs and innovative markers using noninvasive sensors and a home-telemonitoring system. Six HF clinics in Germany and Spain participated in the collection of the clinical data. Patients were included in the study if they had chronic HF with an elevated N-terminal of the prohormone brain natriuretic peptide (NT-proBNP \geq 500 pg/ml), were taking at least 40 mg/day of furosemide or an equivalent, and were in the New York Heart Association (NYHA) functional class II, III, or IV. They were excluded if they had the following: severe chronic obstructive pulmonary disease (COPD GOLD Class > 2), primary pulmonary hypertension, renal insufficiency requiring dialysis, a psychiatric or neurological disorder of moderate to severe degree (eg, dementia, schizophrenia, substance disorder, psychotic depression), prior acute myocardial infarction or coronary artery bypass grafting (CABG) in the previous 3 months. Ethical approval was provided by the Medical Ethics Committees in the 2 respective countries.

Of 148 patients recruited from October 2008 to July 2010, 108 had the system installed and data recorded; 3 did not fit the criteria, 3 were unavailable at installation, 1 died before installation, and 33 withdrew before system installation. Of the remaining 108 users, 17 used the system on less than 30 occasions, leaving 91 patients as the focus of this exploratory analysis. Their mean (SD) age was 63 (12) years and 64 were men. Mean weight was 84 (19) kg, mean BMI was 29 (6) kg/m², and mean left ventricular ejection fraction (LVEF) was 31 (12) %. Most patients had mild (NYHA class II: 60%) or moderate (NYHA class III: 36%) symptoms. Etiology was ischemic in 47%, idiopathic dilated cardiomyopathy in 31%, valvular disease in 5%, and other in 9%. Comorbidities included hypertension (68%), diabetes (37%), atrial fibrillation (36%), renal dysfunction (28%) and COPD (13%). Treatment included angiotensin converting enzyme (ACE) or angiotensin receptor blockers (ARB) (87%), beta-blockers (88%), MRA (53%), diuretics (84%), digoxin (21%), and implantable cardioverter-defibrillator/cardiac resynchronization therapy (ICD/CRT) (23%/14%). The average monitoring time was 10 months, during which 19 patients were hospitalized one or more times due to decompensated HF, with a total of 24 decompensated HF hospitalizations. The adverse events were adjudicated by an advisory committee.

Daily Measurements of Body Weight and NITTI

Patients were instructed on how to perform measurements of body weight and NITTI. Measurements were carried out in the morning before eating breakfast. Body weight was collected using a weight scale (Philips Medical Systems, Andover, Massachusetts, USA), which automatically logged the measurements (accuracy ± 0.1 kg). TTI was measured using a

XSL•FO

wearable bio-impedance vest [28], shown in Figure 1. The vest measures TTI at several electrical frequencies (10 kHz-1MHz). These recordings give a characterization of the electrical properties of the tissue, as described by the Cole-Cole model [29]. At low measurement frequencies, biological tissue impedance is mainly determined by the extracellular fluid content and characteristics. At higher frequencies, electrical properties are determined by both the intracellular and extracellular fluid content. Multi-frequency measurements of

thoracic bio-impedance therefore allow isolation of the Cole parameters that indirectly reflect either the intracellular or extracellular fluid content. We used the *external resistance* derived from the Cole-Cole model, since this indirectly reflects extracellular water, which is the component associated with decompensation. In another study, we have shown that this metric tracks changes in symptoms and fluid loss during treatment for decompensated HF [17].

Figure 1. The bioimpedance vest shown by a model subject correctly applying it across the chest. Textile electrodes on each side of the flexible measurement panel inject currents at different frequencies and register the resulting voltage to calculate the impedance parameter relating to extracellular fluid volume.



Alarm and Event Definition

The weight and NITTI data were applied to published algorithms (detailed description in Multimedia Appendix 1), to predict the onset of decompensation prior to subsequent hospitalization due to worsening heart failure. The output of these algorithms, the *output index*, could be as simple as the difference between the current measurement and the measurement made 2 days previously, or a more complex calculation (eg, one based on cumulative sums). An alert is triggered when the output index exceeds a specific threshold.

The predictive power of the algorithms was assessed by exploring their ability to alarm within a prespecified period before a hospitalization due to worsening heart failure. Changes in NITTI are thought to precede changes in weight prior to hospitalization [12,21]. Depending on the measure used, previous studies have considered alerting periods from 2 weeks [23] up to one month [30] before hospitalization. In this study, a 2-week period was chosen as an adequate period before a hospitalization, during which alarms should be raised, giving time for the patient or clinician to act. Alerts occurring outside of this period were counted as false alarms. Short periods of a few days at the start of monitoring, end of monitoring, and directly following a hospitalization did not fit into any 2-week division and were subsequently removed from the analysis.

Performance Assessment of Algorithms

Three types of alert algorithms are compared in this study: rule-of-thumb (RoT) [21,23,26], moving average convergence divergence (MACD) [23], and cumulative sum control chart (CUSUM) [31]. The qualitative differences between these are shown in Figure 2. Rule of thumb (RoT) methods provide a noisy measure for which chance readings have a large effect, sometimes with no underlying trend; however, they also provide a fast response to changes. Moving averages (MACD) react more slowly but follow underlying trends better, in both directions. Cumulative sums (CUSUM) provide uni-directional detection and lead to longer sustained alerts. For a detailed description of the definitions of each algorithm see Appendix 1. The predictive performance of the algorithms was compared using receiver operator curve (ROC) analysis. The sensitivity and specificity of each algorithm was calculated by dividing the measurement data into periods of 2 weeks, in such a way that a period containing a decompensated hospitalization would end when the hospitalization event occurred. This led to the following definitions:

1. *True positive*: An alarm during the 2 weeks preceding a hospitalization;

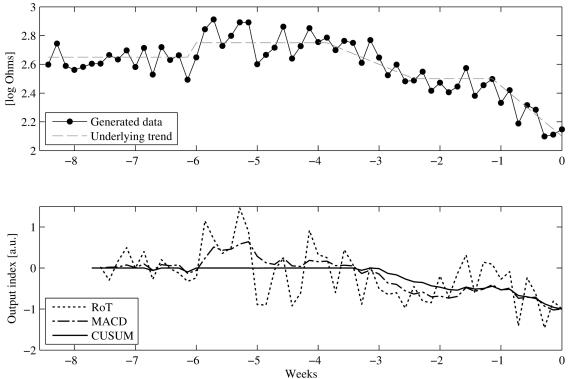
2. False positive: An alarm during any other 2-week period;

3. True negative: A 2-week period without any alarms;

4. *False negative*: A 2-week period ending in a hospitalization without any alarms.



Figure 2. Generated example data with the underlying trend in NITTI are shown in the top graph. The resulting output of the three algorithms, normalized to the last measure to show the qualitative difference between the algorithms, is shown in the bottom graph.



Algorithm Selection and Optimization

Each of the algorithms considered in this study (RoT, MACD, CUSUM) have modifiable parameters that will alter their behavior and ultimately their predictive performance. We tested the performance of each algorithm for a range of possible parameter values. For the RoT algorithms, the number of days (*d*) between the measurements used to calculate the difference was varied from 1 day to 21 days. In the MACD algorithm, the long-term average parameter N_l was varied between 10 and 50 days in increments of 5 days, and the short-term moving average parameter N_s was varied between 1 and 10 days. In the CUSUM algorithm, the parameter determining the length of the running mean and standard deviation (*d*) was varied between 10 and 30 days in increments of 5 days, and the parameter determining the depreciation of the accumulated sum (*c*) was evaluated between 0.5 and 1.5, in increments of 0.2.

Segmentation of the data into 2-week periods results in substantially more periods without an HF-related hospitalization compared to those with one. To avoid producing algorithms that raise a large number of false positive alarms, previous studies have focused only on alarms with high specificity [20,30]. In this investigation, the best parameters were chosen to be those that maximized the area under the curve for thresholds with a specificity >95%. The output index for each algorithm was then normalized to allow the correct estimation of the ROC curves during the cross-validation procedure described below.

Parameter optimization can lead to models that overfit the data, which then would not generalize well to other data sets. To minimize these effects, we implemented a stratified leave-patient-out cross-validation (CV) method for the

```
http://medinform.jmir.org/2016/1/e3/
```

parameters in the RoT, MACD, and CUSUM algorithms. This procedure randomly splits the data into 8 groups, while maintaining the number of patients and decompensation events in each group. The parameters were then optimized for the data with one group left out. The data from the left-out group were then used to evaluate the performance of the optimized parameters. This was repeated until all groups had been left out once. The left-out groups were then recombined to provide an unbiased ROC curve. The optimal threshold for the output index was chosen to be the Youden point with specificity larger than 90%.

Statistics

Comparisons between the recorded measurements and the output index for the different algorithms in the 2 weeks preceding hospitalization and all other periods were tested with a mixed-effect model using patient specific intercepts as random effects. An arbitrary significance of 0.05 was assumed throughout. Missing data due to adherence issues were removed from the analysis by excluding periods in which less than 3 [32] measurements per week were found. In the case of algorithms that needed previous data points to estimate trends, a linear imputation between adjacent data points was carried out. It should be noted that when the algorithms processed the data, imputations were only made on data that would have been available for a system running in real time; no imputations using future values were done. NITTI measurements were log-transformed to adjust for skewness. All listed algorithms were developed and evaluated using the software suite MATLAB 7.13.0.564.

Results

Data Characteristics

Among the 91 patients for whom data were included in the analysis, 24 heart failure-related hospitalizations occurred in 19 patients. Of the 24 hospitalizations, 9 had less than 3 weekly weight recordings and 12 had less than 3 weekly impedance recordings preceding the hospitalization, and were excluded from the analysis. The minimum window for the CUSUM algorithm excluded an additional 2 for its analysis.

Prediction Performance

The predictive performance of guideline-based rules and published algorithms using weight are presented in Table 1. With the exception of those rules with very low specificity (ie, <60%), all rules based on short-term increases had low sensitivity when applied to the data (typically <25%). Rules

based on longer-term increases showed higher sensitivity; however, only one had a specificity >90%. The MACD algorithm with the parameter proposed by Zhang et al. [23] outperformed the other weight algorithms.

The cross-validation analyses of the developed models based on published algorithms are presented in Figure 3. The RoT-based algorithms using weight have poor sensitivity at a specificity between 90-100%, with performance close to random chance. This poor sensitivity was also observed when evaluating previous published guidelines using windows between 2 and 3 days (Table 1). As expected, this sensitivity increased when longer windows and/or lower thresholds were used, but at the cost of a lower specificity.

The MACD algorithm improved performance for both weight and impedance. The CUSUM algorithm improved performance for NITTI. The performance of trend algorithms was superior to previously published algorithms (Table 1).

Table 1. Performance of different weight algorithms in anticipating an upcoming decompensation.

Source	Weight algorithm	Sensitivity Specificity		PPV ^a	NPV ^b	
		%	%	%	%	
Guideline issuing bodies	>2 lbs ^c in 1 day [10]	67	56	1.4	99.5	
	>2 kg in 3 days [9]	13	87	0.9	99.1	
	>4 lbs ^c in 1 week [10]	27	87	1.8	99.2	
Existing literature	Random chance	50	50	0.9	99.1	
	>2 lbs in 1 day or >3 lbs in 3 days [26]	in 3 days [26] 73 50	1.3	99.5		
	>2 lbs in 1 day or >5 lbs in 3 days [26]	67	56	1.4	99.4	
	>3 lbs in 1 day or >5 lbs in 3 days [26]	5 lbs in 3 days [26] 13 82	82	0.7	99.1	
	>3 lbs in 1 day or >7 lbs in 3 days [26]	7	83	0.4	99.0	
	>4 lbs in 1 day or >7 lbs in 3 days [26]	7	93	0.9	99.1	
	>4 lbs in 1 day or >9 lbs in 3 days [26]	7	93	0.9	99.1	
	>5 lbs in 1 day or >9 lbs in 3 days [26]	0	100	_	99.1	
	>2 lbs in 1 week [21]	80	45	1.3	99.6	
	>5 lbs in 1 week [21]	20	94	2.7	99.2	
	>4 lbs in a 5 to 80 days MACD ^d [23]	20	97	6.3	99.3	

^aPPV: positive predictive value

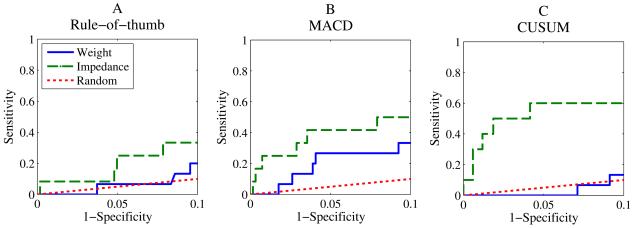
^bNPV: negative predictive value

^cTo convert to kilograms multiply by 0.45

^dMACD: moving average convergence divergence



Figure 3. ROC curves from the cross-validated evaluation for the three considered algorithms in the specificity range from 0.9 to 1. A shows the rule of thumb algorithm, B the MACD algorithm, and C the CUSUM algorithm. Performance using NITTI measures is shown with the dashed green line, weight is shown with the blue line, and random chance is portrayed by the red dotted line.



Optimal Parameters

The output of the 2 best performing algorithms for weight and impedance with optimal parameters (maximum Youden index with specificity >90%) is shown in Figure 4. Clear trends in both weight and impedance can be seen for Patient 1 and both algorithms managed to alert before the decompensation; a full week in advance for impedance and a day in advance for weight.

Patient 2, on the other hand, had no or weakly visible trends, which were not enough to trigger an alert. The patient did exhibit large daily weight fluctuations, which could have indicated instability; however, this was not picked up by the algorithms. The optimal parameters for all 3 algorithms for weight and impedance are shown in Table 2, together with the cross-validated performance measures. Both trend algorithms using NITTI outperformed the weight algorithms.

Table 2. Cross-validated performance measures of the algorithms at the maximum Youden index within a specificity of 90-100%.

Optimal algorithms ^a		Sensitivity %	Specificity %	PPV ^b %	NPV ^c %
Weight					
	RoT ^d : >2.7 kg in 17 days	20	90	1.95	99.2
	MACD ^e : >0.62 kg (N _s =9, N _l = 20 days)	33	91	3.2	99.3
	CUSUM ^f :>8.7 with 10-day average, c=0.75	13	91	1.4	99.1
NITTI ^g					
	RoT: <-0.27 (log ohm) in 21 days	33	92	4.2	99.2
	MACD: <-0.059 (log ohm) (N _s =9, N _l = 35 days)	50	92	5.9	99.5
	CUSUM: <-7.8 with 20-day average, c=0.75	60	96	10.9	99.6

^aThe optimal parameters and thresholds were estimated from the full data (for stability and variance of cross-validated parameters and thresholds, see Table 3).

^bPPV: positive predictive value

^cNPV: negative predictive value

^dRoT: rule of thumb

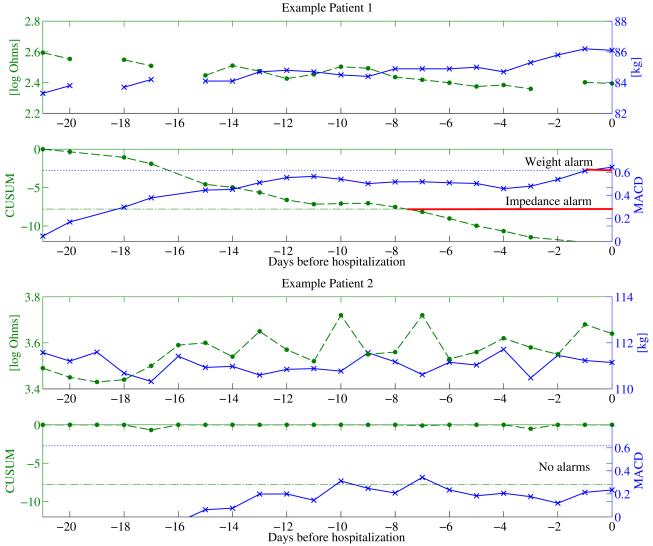
^eMACD: moving average convergence divergence

^fCUSUM: cumulative sums

^gNITTI: noninvasive transthoracic bio-impedance



Figure 4. Three weeks of telemonitoring data from two patients with high compliance before an upcoming decompensation. Circles correspond to NITTI measurements and the NITTI-CUSUM algorithm and crosses correspond to weight measurements and the weight-MACD algorithm. Optimal thresholds are shown as dash-dotted lines in green for NITTI and dotted blue lines for weight.



Algorithmic Stability

The use of a cross-validation procedure to minimize biased performance measures generated several plausible parameters for the tested algorithms; these are presented in Table 3. In general, RoT had lower variance in estimated parameters than MACD, which in turn had lower variance than CUSUM, coinciding with the increasing complexity of the algorithms. Parameter variance was especially high for the weight CUSUM algorithm, which could explain the poor performances when compared to MACD.

Mean values for weight, impedance, and the respective output indices of the optimal algorithms during periods preceding a hospitalization compared to the other periods are shown in Table 4. A statistically significant difference was only found for the NITTI measurements and algorithms based upon NITTI.



Cuba Gyllensten et al

Table 3. Mean, standard deviation, and individual values for the estimated optimal parameters in each of the 8 folds created using the described stratified cross-validation procedure.

Measure	Body weight							Transthoracic impedance									
CV ^a step		1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
RoT ^b																	
	Threshold	3.5 (0.08)			-0.31 (0.035)												
		3.5	3.56	3.4	3.56	3.45	3.6	3.4	3.4	-0.3	-0.31	-0.3	-0.3	-0.3	-0.3	-0.3	-0.4
	Days				14.4	4 (3.7)							20.	5 (1.41)		
		11	17	11	17	17	20	11	11	21	17	21	21	21	21	21	21
MACD ^c																	
	Threshold				0.8	(0.38)							-0.1	0 (0.01	4)		
		1.59	0.62	0.31	0.62	0.62	0.97	0.62	0.95	-0.12	-0.1	-0.1	-0.1	-0.1	-0.09	-0.09	-0.1
	Short-term avg. window				8.6	(1.19)							8.	1 (0.99)		
		8	9	8	9	9	10	9	9	9	8	8	8	8	9	9	6
	Long-term avg. window				25.6	(10.84	•)						36.	.3 (3.54	l)		
		50	20	15	20	25	30	20	25	45	35	35	35	35	35	35	35
CUSUM ^d																	
	Threshold				11.0	(7.87)						-8.1	13 (2.6	5)		
		30	8.7	8.7	8.7	6.9	8.1	8.7	8.1	-7.8	-10.3	-7.8	-7.8	-11.1	-4.40	-11.14	-4.6
	Days				26.9	(18.3)						18.	.8 (2.31)		
		50	10	10	10	40	40	10	45	20	20	20	20	15	20	15	20
	Depreciation				1.13	(0.40)			0.75 (0.19)							
		1.5	0.75	0.75	0.75	1.5	1.5	0.75	1.5	0.75	0.75	0.75	0.75	0.50	1	0.50	1

^aCV: Cross-validation

^bRoT: rule of thumb

^cMACD: moving average convergence divergence

^dCUSUM: cumulative sums

Table 4. Population mean output index values for RoT, MACD, and CUSUM algorithms using the optimal parameters (see 2) in the 2-week period preceding a hospitalization compared to all other periods.

Measure	Mean (SD) value in 2-week period be- fore decompensation	Mean (SD) value in nondecompensa- tion periods	Statistical significance ^d
Weight (kg)	83 (10)	84 (19)	.97
Weight-RoT ^a (kg)	0.3 (1.2)	0.06 (0.87)	.76
Weight-MACD ^b (kg)	0.08 (0.30)	0.02 (0.22)	.24
Weight-CUSUM ^c (kg)	1.9 (2.7)	0.8 (1.3)	.58
TTI (log Ohm)	3.0 (0.3)	3.4 (0.3)	<.001
TTI-RoT (log Ohm) ^a	-0.07 (0.12)	0.00 (0.08)	<.001
TTI-MACD (log Ohm) ^a	-0.032 (0.044)	0.003 (0.028)	<.001
TTI-CUSUM (log Ohm) ^a	-6.4 (9.4)	-0.7 (2.0)	<.001

^aRoT: rule of thumb

^bMACD: moving average convergence divergence

^cCUSUM: cumulative sums

^dEstimated with a mixed-effect model with patient specific random effects. For the algorithms the cross-validation output was used.

http://medinform.jmir.org/2016/1/e3/

Cuba Gyllensten et al

Discussion

Principal Findings

The main finding of the present study is that change in NITTI is a stronger predictor of an impending decompensation compared to changes in weight (cross-validation estimate was 60% for NITTI-CUSUM vs 33% for Weight-MACD) and that both measurements benefit from trend detection algorithms. Mean values of NITTI in the 2-week period preceding a decompensation event were lower than in nondecompensation periods (P<.001).

Fluid overload is one of the leading causes for HF hospitalization and body weight increase has been linked to an increased risk of hospitalization [21]. However, directly applying a weight gain difference to predict imminent decompensation is challenging. This study corroborates the findings of Zhang [23] and Abraham [26], who also reported low predictive ability of alarms using short-term weight change. Short-term weight increase will detect a large and rapid fluid accumulation. Our evaluation of the rule suggested by the ESC guidelines is that it has high specificity but it is not a very sensitive method to predict HF hospitalization, as gradual weight increases are missed. A moving average algorithm focuses on progressive changes in weight, removing much of the inherent variability in weight measurements and errors due to the home setting in which patients might deviate from the measurement protocol, and daily changes due to dietary and fluid intake are averaged out. This could explain why lower threshold values led to higher sensitivity while still retaining specificity.

The increase in thoracic fluid due to congestion should decrease impedance measurements. Several studies have reported positive results from algorithms using impedance to detect decompensations [19,20,33]. To test algorithms proposed for decompensation detection using impedance measurements, we employed a cross-validation procedure to estimate performances. The results are similar, although on the lower side of what has been reported for ITI in terms of sensitivity (76.4% [26], 76.9% [33], 60% [34]), perhaps partly accounted for by the robust methods we employed. Reported performances from feasibility studies usually decline in later prospective studies [35], which the leave-subject-out protocol is designed to emulate.

Comparisons between predicted performances of weight and impedance measurements in Figure 3 show that impedance is the stronger predictor. This is also suggested by the analyses of the mean output index in the 2 weeks preceding a decompensation (Table 4), for which a statistical difference was found compared to periods without decompensation for all impedance algorithms as well as the impedance value, but not for any of the weight algorithms. Abraham et al. [26] also showed a higher sensitivity for impedance measurements when compared to weight. However, we showed that the gap in performance could be made smaller with more sophisticated weight trend algorithms compared to the rules suggested by Abraham (in which the 3 rules with a specificity >90% had a maximum sensitivity of 7%). Sensitivity to fluid build-up in the lungs, whether through redistribution of fluids or retention, could explain the increased performance of impedance when

```
http://medinform.jmir.org/2016/1/e3/
```

compared to body weight [11]. Similarly, weight loss from malnutrition might mask fluid accumulation in weight measurements, which would still be picked up by NITTI. The focus in this study on high specificity algorithms might also have put weight algorithms at a slight disadvantage; evidence of this can be found in the stability analysis (Table 3), in which the high parameter variance for the weight-CUSUM algorithm could have resulted from the difficulty of finding a highly specific algorithm, which led to a negative impact on its cross-validated performance.

The difficulty in assessing prediction algorithms is known [36]. Different evaluation metrics can show diverging results, because they shed light on different aspects of performance. Definitions of what constitutes a true positive and false positive have a great effect on performance. In this study, we focused on algorithms with high specificity evaluated using 2-week intervals, with the best-performing alarm having a sensitivity rate of 60%. Although this catches several patients at a high specificity, it still raises unexplained alarms and has a relatively low positive predictive value of 10.9% for impedance and 3.2% for weight. A measure focusing on the workload associated with managing these alerts, such as false alarms per patient year has been used by several other studies as a surrogate specificity metric [26,33-35]. Defined as an alert not resulting in a hospitalization, the NITTI-CUSUM algorithm has a cross-validated estimate of 0.48 false alarms per patient year. These seemingly contradictory performance measures can be explained by the rarity of 2-week periods resulting in hospitalization, when compared to the full amount of telemonitoring data. An alarm that goes on for 5 weeks would cross three 2-week periods and could generate 3 false positives; however, using the false-alarm metric it would only add one false alarm.

Therefore, the positive predictive value of 10.9% should be seen in the context of 2-week windows having both high specificity and sensitivity and compared to the relatively low predictive value of current weight algorithms.

Low levels of positive predictive value have also been observed in many other studies evaluating prediction algorithms from daily measurements [35,37,38]. The concept of predicting future events might be less realistic than providing indications that could be acted upon. This approach could tailor actions depending on which monitored sign was detected. Indeed, many signs that have been linked to deterioration, for example, arrhythmias [39], breathing rates [38], and heart-rate variability [40], can be detected noninvasively and may be included in such an approach. Importantly, the implementation of better decompensation algorithms will reduce the number of clinical alerts that would need to be dealt with by a telehealth nurse or physician. This will result in better resource utilization, with the management of larger patient caseloads and, therefore, a reduction in the costs of patient management.

Limitations

Although clinicians were blinded to the observational data, they could have intervened based on increased weight data for worsening patients. If such interventions did not result in a hospitalization, they were not recorded in this study and might have negatively affected the results. In the SENSE-HF trial [37],

XSL•FO RenderX

a substantial increase in positive predictive value was reported after including signs and symptoms of worsening HF diagnosed by a physician rather than only adjudicated HF hospitalizations; therefore, it could also be expected that several false positives were due to "mild" decompensations. Indeed, it is possible that patients often self-correct decompensation by reducing dietary salt, increasing adherence to medication, or even by taking extra doses of diuretic. Changes in environmental temperature might also affect compensation. In this study, high specificity alerts were explored. However, sacrificing specificity for improved sensitivity may be a good complement if management of alerts can be handled by patients without resorting to professional advice. Combining specific alerts with a strategy of health maintenance might be superior to one of only crisis detection and management [41]. Most patients are interested and able to contribute to their care if they are given the information and confidence to do so. Remote monitoring provides a safe environment or safety net to encourage such behavior.

Incorrectly using the measurement equipment could have caused erroneous values with the net effect of lowered performances. The surface on which the scales sit, their accuracy, clothing, and use by other family members can all cause problems with measurement. Bio-impedance weight scales (a different technology from NITTI) require patients to remove their socks and shoes and hence may improve the consistency of measurement. Giving patients feedback and asking them to recheck their weight if it falls out of the expected range are all likely to improve the data quality on which the algorithms are based. The limited amount of data available for this study makes generalizations difficult. Application of cross-validation procedures were employed to minimize this effect; however, the calculated percentage values were ultimately derived from a small set of subjects and should therefore be seen as qualitative indicators of performance.

Conclusion

Daily measurements of transthoracic impedance using a vest with textile electrodes is a feasible way to monitor HF and provides a more accurate indication of upcoming decompensations when compared to weight for all 3 algorithms tested (RoT, MACD, and CUSUM). Trend detection algorithms outperformed RoT measures suggesting that tracking the progression is more important than direct measures of change, which currently are suggested by guidelines.

However, the low positive predictive value of all the algorithms tested did not allow accurate prediction of impending HF hospitalizations. Implementation of trend detection algorithms might better serve as indications of worsening, which, when integrated with other clinical measures, could be useful for treatment management. The promising results from this investigation warrant further trials with noninvasive TTI as a technology for the management of HF, perhaps connected to actionable alerts. These alerts would promote a strategy of "health maintenance" to keep the patient as close to their ideal state as possible on a daily basis, which could be combined with a strategy of "crisis detection and management" if the first strategy failed.

Acknowledgments

This work was supported by the EU Marie Curie Network iCareNet under grant number 264738. Data were provided by the MyHeart project, which was partially financed by the EU FP6 program under grant number 507816.

Conflicts of Interest

ICG is a PhD student employed at Philips Research. AGB, HR, and JH are employed by Philips Research. JGFC and KGM have received departmental research support from Philips.

Multimedia Appendix 1

Detailed description of algorithms to detect decompensated HF.

[PDF File (Adobe PDF File), 128KB-Multimedia Appendix 1]

References

- 1. Mosterd A, Hoes AW. Clinical epidemiology of heart failure. Heart 2007 Sep;93(9):1137-1146 [FREE Full text] [doi: 10.1136/hrt.2003.025270] [Medline: 17699180]
- 2. Berry C, Murdoch DR, McMurray JJ. Economics of chronic heart failure. Eur J Heart Fail 2001 Jun;3(3):283-291. [Medline: 11377998]
- Maggioni AP, Dahlström U, Filippatos G, Chioncel O, Crespo LM, Drozdz J, Heart Failure Association of the European Society of Cardiology (HFA). EURObservational Research Programme: Regional differences and 1-year follow-up results of the Heart Failure Pilot Survey (ESC-HF Pilot). Eur J Heart Fail 2013 Jul;15(7):808-817. [doi: <u>10.1093/eurjhf/hft050</u>] [Medline: <u>23537547</u>]
- Cleland JGF, McDonagh T, Rigby AS, Yassin A, Whittaker T, Dargie HJ, National Heart Failure Audit Team for England and Wales. The national heart failure audit for England and Wales 2008-2009. Heart 2011 Jun;97(11):876-886. [doi: 10.1136/hrt.2010.209171] [Medline: 21173198]

RenderX

- Giamouzis G, Kalogeropoulos A, Georgiopoulou V, Laskar S, Smith AL, Dunbar S, et al. Hospitalization epidemic in patients with heart failure: Risk factors, risk prediction, knowledge gaps, and future directions. J Card Fail 2011 Jan;17(1):54-75. [doi: 10.1016/j.cardfail.2010.08.010] [Medline: 21187265]
- Inglis SC, Clark RA, McAlister FA, Stewart S, Cleland JGF. Which components of heart failure programmes are effective? A systematic review and meta-analysis of the outcomes of structured telephone support or telemonitoring as the primary component of chronic heart failure management in 8323 patients: Abridged Cochrane Review. Eur J Heart Fail 2011 Sep;13(9):1028-1040. [doi: 10.1093/eurjhf/hfr039] [Medline: 21733889]
- Desai AS, Stevenson LW. Connecting the circle from home to heart-failure disease management. N Engl J Med 2010 Dec 9;363(24):2364-2367. [doi: 10.1056/NEJMe1011769] [Medline: 21080836]
- 8. Boriani G, Da Costa A, Ricci RP, Quesada A, Favale S, Iacopino S, MORE-CARE Investigators. The MOnitoring Resynchronization dEvices and CARdiac patiEnts (MORE-CARE) randomized controlled trial: Phase 1 results on dynamics of early intervention with remote monitoring. J Med Internet Res 2013;15(8):e167 [FREE Full text] [doi: 10.2196/jmir.2608] [Medline: 23965236]
- 9. McMurray JJV, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, ESC Committee for Practice Guidelines. ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur Heart J 2012 Jul;33(14):1787-1847 [FREE Full text] [doi: 10.1093/eurheartj/ehs104] [Medline: 22611136]
- 10. Heart Failure Society of America. Module 4: Self Care Following Your Treatment Plan and Dealing with your Symptoms Internet URL: <u>http://www.hfsa.org/hfsa-wp/wp/module-4/</u> [accessed 2015-06-18] [WebCite Cache ID 6ZNZ4K4RY]
- 11. Cotter G, Felker GM, Adams KF, Milo-Cotter O, O'Connor CM. The pathophysiology of acute heart failure--is it all about fluid accumulation? Am Heart J 2008 Jan;155(1):9-18. [doi: <u>10.1016/j.ahj.2006.02.038</u>] [Medline: <u>18082483</u>]
- 12. Gheorghiade M, Follath F, Ponikowski P, Barsuk JH, Blair JEA, Cleland JG, European Society of Cardiology, European Society of Intensive Care Medicine. Assessing and grading congestion in acute heart failure: A scientific statement from the acute heart failure committee of the heart failure association of the European Society of Cardiology and endorsed by the European Society of Intensive Care Medicine. Eur J Heart Fail 2010 May;12(5):423-433 [FREE Full text] [doi: 10.1093/eurjhf/hfq045] [Medline: 20354029]
- Abraham WT, Adamson PB, Bourge RC, Aaron MF, Costanzo MR, Stevenson LW, CHAMPION Trial Study Group. Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: A randomised controlled trial. Lancet 2011 Feb 19;377(9766):658-666. [doi: 10.1016/S0140-6736(11)60101-3] [Medline: 21315441]
- Vanderheyden M, Houben R, Verstreken S, Ståhlberg M, Reiters P, Kessels R, et al. Continuous monitoring of intrathoracic impedance and right ventricular pressures in patients with heart failure. Circ Heart Fail 2010 May;3(3):370-377 [FREE Full text] [doi: 10.1161/CIRCHEARTFAILURE.109.867549] [Medline: 20197559]
- 15. Packer M, Abraham WT, Mehra MR, Yancy CW, Lawless CE, Mitchell JE, Prospective Evaluation and Identification of Cardiac Decompensation by ICG Test (PREDICT) Study Investigators and Coordinators. Utility of impedance cardiography for the identification of short-term risk of clinical decompensation in stable patients with chronic heart failure. J Am Coll Cardiol 2006 Jun 6;47(11):2245-2252 [FREE Full text] [doi: 10.1016/j.jacc.2005.12.071] [Medline: 16750691]
- Malfatto G, Villani A, Rosa FD, Rella V, Oldani M, Giglio A, et al. Correlation between trans and intra-thoracic impedance and conductance in patients with chronic heart failure. J Cardiovasc Med (Hagerstown) 2014 Sep 15. [doi: <u>10.2459/JCM.000000000000177</u>] [Medline: <u>25226339</u>]
- Cuba-Gyllensten I, Gastelurrutia P, Riistama J, Aarts R, Nuñez J, Lupon J, et al. A novel wearable vest for tracking pulmonary congestion in acutely decompensated heart failure. Int J Cardiol 2014 Nov 15;177(1):199-201. [doi: 10.1016/j.ijcard.2014.09.041] [Medline: 25499378]
- Amir O, Rappaport D, Zafrir B, Abraham WT. A novel approach to monitoring pulmonary congestion in heart failure: Initial animal and clinical experiences using remote dielectric sensing technology. Congest Heart Fail 2013;19(3):149-155 [FREE Full text] [doi: 10.1111/chf.12021] [Medline: 23350643]
- 19. Shochat M, Shotan A, Kazatsker M, Gurovich V, Shochat I, Naiman E, et al. Lung impedance monitoring in the outpatient clinic predicts hospitalizations of patients with decompensated heart failure and enables early therapy to prevent hospitalizations. In: J Am Coll Cardiol. 2011 Apr Presented at: American College of Cardiology (ACC) 60th Annual Scientific Session; April 2 5, 2011; New Orleans p. E1259.
- 20. Anand IS, Wilson Tang WH, Greenberg BH, Chakravarthy N, Libbus I, Katra RP, Music Investigators. Design and performance of a multisensor heart failure monitoring algorithm: Results from the multisensor monitoring in congestive heart failure (MUSIC) study. J Card Fail 2012 Apr;18(4):289-295. [doi: 10.1016/j.cardfail.2012.01.009] [Medline: 22464769]
- Chaudhry SI, Wang Y, Concato J, Gill TM, Krumholz HM. Patterns of weight change preceding hospitalization for heart failure. Circulation 2007 Oct 2;116(14):1549-1554 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.107.690768] [Medline: 17846286]
- 22. Whellan DJ, Ousdigian KT, Al-Khatib SM, Pu W, Sarkar S, Porter CB, PARTNERS Study Investigators. Combined heart failure device diagnostics identify patients at higher risk of subsequent heart failure hospitalizations: results from PARTNERS HF (Program to Access and Review Trending Information and Evaluate Correlation to Symptoms in Patients With Heart

RenderX

Failure) study. J Am Coll Cardiol 2010 Apr 27;55(17):1803-1810 [FREE Full text] [doi: 10.1016/j.jacc.2009.11.089] [Medline: 20413029]

- Zhang J, Goode KM, Cuddihy PE, Cleland JGF, TEN-HMS Investigators. Predicting hospitalization due to worsening heart failure using daily weight measurement: Analysis of the Trans-European Network-Home-Care Management System (TEN-HMS) study. Eur J Heart Fail 2009 Apr;11(4):420-427 [FREE Full text] [doi: 10.1093/eurjhf/hfp033] [Medline: 19252210]
- 24. Lewin J, Ledwidge M, O'Loughlin C, McNally C, McDonald K. Clinical deterioration in established heart failure: What is the value of BNP and weight gain in aiding diagnosis? Eur J Heart Fail 2005 Oct;7(6):953-957 [FREE Full text] [doi: 10.1016/j.ejheart.2005.06.003] [Medline: 16227134]
- 25. Ledwidge MT, O'Hanlon R, Lalor L, Travers B, Edwards N, Kelly D, et al. Can individualized weight monitoring using the HeartPhone algorithm improve sensitivity for clinical deterioration of heart failure? Eur J Heart Fail 2013 Apr;15(4):447-455 [FREE Full text] [doi: 10.1093/eurjhf/hfs186] [Medline: 23204211]
- Abraham WT, Compton S, Haas G, Foreman B, Canby RC, Fishel R, FAST Study Investigators. Intrathoracic impedance vs daily weight monitoring for predicting worsening heart failure events: Results of the Fluid Accumulation Status Trial (FAST). Congest Heart Fail 2011;17(2):51-55 [FREE Full text] [doi: 10.1111/j.1751-7133.2011.00220.x] [Medline: 21449992]
- 27. Habetha J. The MyHeart project--fighting cardiovascular diseases by prevention and early diagnosis. In: Conf Proc IEEE Eng Med Biol Soc. 2006 Aug Presented at: 28th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; 31 Aug 3 Sep; New York p. 6746-6749. [doi: 10.1109/IEMBS.2006.260937]
- Reiter H, Muehlsteff J, Sipilä A. Medical application and clinical validation for reliable and trustworthy physiological monitoring using functional textiles: Experience from the HeartCycle and MyHeart project. In: Conf Proc IEEE Eng Med Biol Soc. 2011 Presented at: 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society; Aug 30 - Sep 3; Boston p. 3270-3273. [doi: 10.1109/IEMBS.2011.6090888]
- 29. Gabriel S, Lau RW, Gabriel C. The dielectric properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. Phys Med Biol 1996 Nov;41(11):2251-2269. [Medline: <u>8938025</u>]
- Sack S, Wende CM, Nägele H, Katz A, Bauer WR, Barr CS, et al. Potential value of automated daily screening of cardiac resynchronization therapy defibrillator diagnostics for prediction of major cardiovascular events: Results from Home-CARE (Home Monitoring in Cardiac Resynchronization Therapy) study. Eur J Heart Fail 2011 Sep;13(9):1019-1027 [FREE Full text] [doi: 10.1093/eurjhf/hfr089] [Medline: 21852311]
- Adamson PB, Zile MR, Cho YK, Bennett TD, Bourge RC, Aaron MF, et al. Hemodynamic factors associated with acute decompensated heart failure: Part 2--use in automated detection. J Card Fail 2011 May;17(5):366-373. [doi: 10.1016/j.cardfail.2011.01.011] [Medline: 21549292]
- 32. Lyngå P, Persson H, Hägg-Martinell A, Hägglund E, Hagerman I, Langius-Eklöf A, et al. Weight monitoring in patients with severe heart failure (WISH). A randomized controlled trial. Eur J Heart Fail 2012 Apr;14(4):438-444 [FREE Full text] [doi: 10.1093/eurjhf/hfs023] [Medline: 22371525]
- 33. Yu C, Wang L, Chau E, Chan RH, Kong S, Tang M, et al. Intrathoracic impedance monitoring in patients with heart failure: Correlation with fluid status and feasibility of early warning preceding hospitalization. Circulation 2005 Aug 9;112(6):841-848 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.104.492207] [Medline: 16061743]
- 34. Vollmann D, Nägele H, Schauerte P, Wiegand U, Butter C, Zanotto G, Hill Michael R S, European InSync Sentry Observational Study Investigators. Clinical utility of intrathoracic impedance monitoring to alert patients with an implanted device of deteriorating chronic heart failure. Eur Heart J 2007 Aug;28(15):1835-1840 [FREE Full text] [doi: 10.1093/eurheartj/ehl506] [Medline: 17309902]
- 35. Heist EK, Herre JM, Binkley PF, Van Bakel AB, Porterfield JG, Porterfield LM, DEFEAT-PE Study Investigators. Analysis of different device-based intrathoracic impedance vectors for detection of heart failure events (from the Detect Fluid Early from Intrathoracic Impedance Monitoring study). Am J Cardiol 2014 Oct 15;114(8):1249-1256. [doi: 10.1016/j.amjcard.2014.07.048] [Medline: 25150135]
- 36. Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: The long and winding road. Brain 2007 Feb;130(Pt 2):314-333 [FREE Full text] [doi: 10.1093/brain/awl241] [Medline: 17008335]
- 37. Conraads VM, Tavazzi L, Santini M, Oliva F, Gerritse B, Yu C, et al. Sensitivity and positive predictive value of implantable intrathoracic impedance monitoring as a predictor of heart failure hospitalizations: The SENSE-HF trial. Eur Heart J 2011 Sep;32(18):2266-2273 [FREE Full text] [doi: 10.1093/eurheartj/ehr050] [Medline: 21362703]
- 38. Auricchio A, Gold MR, Brugada J, Nölker G, Arunasalam S, Leclercq C, et al. Long-term effectiveness of the combined minute ventilation and patient activity sensors as predictor of heart failure events in patients treated with cardiac resynchronization therapy: Results of the Clinical Evaluation of the Physiological Diagnosis Function in the PARADYM CRT device Trial (CLEPSYDRA) study. Eur J Heart Fail 2014 Jun;16(6):663-670 [FREE Full text] [doi: 10.1002/ejhf.79] [Medline: 24639140]
- 39. Cowie MR, Sarkar S, Koehler J, Whellan DJ, Crossley GH, Wilson Tang WH, et al. Development and validation of an integrated diagnostic algorithm derived from parameters monitored in implantable devices for identifying patients at risk

RenderX

for heart failure hospitalization in an ambulatory setting. Eur Heart J 2013 Aug;34(31):2472-2480 [FREE Full text] [doi: 10.1093/eurheartj/eht083] [Medline: 23513212]

- 40. Adamson PB, Smith AL, Abraham WT, Kleckner KJ, Stadler RW, Shih A, InSync III Model 8042 and Attain OTW Lead Model 4193 Clinical Trial Investigators. Continuous autonomic assessment in patients with symptomatic heart failure: Prognostic value of heart rate variability measured by an implanted cardiac resynchronization device. Circulation 2004 Oct 19;110(16):2389-2394 [FREE Full text] [doi: 10.1161/01.CIR.0000139841.42454.78] [Medline: 15313946]
- 41. Cleland JGF, Antony R. It makes SENSE to take a safer road. Eur Heart J 2011 Sep;32(18):2225-2227 [FREE Full text] [doi: 10.1093/eurheartj/ehr120] [Medline: 21551157]

Abbreviations

ACE: angiotensin converting enzyme ARB: angiotensin receptor blockers CABG: coronary artery bypass grafting **CUSUM:** cumulative sums **CV:** cross-validation **ESC:** European Society of Cardiology HF: Heart Failure HFSA: Heart Failure Association of America ICD/CRT: implantable cardioverter-defibrillator/cardiac resynchronization therapy **ITI:** intrathoracic impedance LVEF: left ventricular ejection fraction MACD: moving average convergence divergence NYHA: New York Heart Association NITTI: noninvasive transthoracic bio-impedance **NPV:** negative predictive value **PAP:** pulmonary arterial pressure **PPV:** positive predictive value **ROC:** receiver operator curve RoT: rule-of-thumb

Edited by G Eysenbach; submitted 18.06.15; peer-reviewed by G Malfatto, K Kim; comments to author 29.07.15; revised version received 09.09.15; accepted 07.10.15; published 18.02.16

<u>Please cite as:</u> Cuba Gyllensten I, Bonomi AG, Goode KM, Reiter H, Habetha J, Amft O, Cleland JGF Early Indication of Decompensated Heart Failure in Patients on Home-Telemonitoring: A Comparison of Prediction Algorithms Based on Daily Weight and Noninvasive Transthoracic Bio-impedance JMIR Med Inform 2016;4(1):e3 URL: <u>http://medinform.jmir.org/2016/1/e3/</u> doi: <u>10.2196/medinform.4842</u> PMID: <u>26892844</u>

©Illapha Cuba Gyllensten, Alberto G Bonomi, Kevin M Goode, Harald Reiter, Joerg Habetha, Oliver Amft, John GF Cleland. Originally published in JMIR Medical Informatics (http://medinform.jmir.org), 18.02.2016. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Medical Informatics, is properly cited. The complete bibliographic information, a link to the original publication on http://medinform.jmir.org/, as well as this copyright and license information must be included.

