Beat-to-beat systolic time-interval measurement from heart sounds and ECG

2012 Physiol. Meas. 33 177

(http://iopscience.iop.org/0967-3334/33/2/177)

IOP Publishing Physiological Measurement

Physiol. Meas. **33** (2012) 177–194 doi:10.1088/0967-3334/33/2/177

Beat-to-beat systolic time-interval measurement from heart sounds and ECG

R P Paiva[1], P Carvalho[1], R Couceiro[1], J Henriques[1], M Antunes[2], I Quintal[3] and J Muehlsteff[4]

[1] Center for Informatics and Systems of the University of Coimbra, Pólo II, 3030-290 Coimbra, Portugal
[2] Cardiothoracic Surgery Center, Hospitals of the University of Coimbra, Praceta Mota Pinto, 3049 Coimbra, Portugal
[3] Hospital Center of Coimbra, Quinta dos Vales, 3041-801 Coimbra, Portugal
[4] Philips Research Laboratories Europe, HTC, 5656AE Eindhoven, The Netherlands

E-mail: ruipedro@dei.uc.pt

Received 8 June 2011, accepted for publication 24 November 2011
Published 19 January 2012
Online at stacks.iop.org/PM/33/177

Abstract

Systolic time intervals are highly correlated to fundamental cardiac functions. Several studies have shown that these measurements have significant diagnostic and prognostic value in heart failure condition and are adequate for long-term patient follow-up and disease management. In this paper, we investigate the feasibility of using heart sound (HS) to accurately measure the opening and closing moments of the aortic heart valve. These moments are crucial to define the main systolic timings of the heart cycle, i.e. pre-ejection period (PEP) and left ventricular ejection time (LVET). We introduce an algorithm for automatic extraction of PEP and LVET using HS and electrocardiogram. PEP is estimated with a Bayesian approach using the signal's instantaneous amplitude and patient-specific time intervals between atrio-ventricular valve closure and aortic valve opening. As for LVET, since the aortic valve closure corresponds to the start of the S2 HS component, we base LVET estimation on the detection of the S2 onset. A comparative assessment of the main systolic time intervals is performed using synchronous signal acquisitions of the current gold standard in cardiac time-interval measurement, i.e. echocardiography, and HS. The algorithms were evaluated on a healthy population, as well as on a group of subjects with different cardiovascular diseases (CVD). In the healthy group, from a set of 942 heartbeats, the proposed algorithm achieved 7.66 $\pm$ 5.92 ms absolute PEP estimation error. For LVET, the absolute estimation error was 11.39 $\pm$ 8.98 ms. For the CVD population, 404 beats were used, leading to 11.86 $\pm$ 8.30 and 17.51 $\pm$ 17.21 ms absolute PEP and

LVET errors, respectively. The results achieved in this study suggest that HS can be used to accurately estimate LVET and PEP.

Keywords: systolic time intervals, cardiac function, heart sound segmentation

(Some figures may appear in colour only in the online journal)

1. Introduction

Since Robert Hooke's discovery of the diagnostic potential of heart sound (HS), cardiac auscultation has been a key instrument for non-invasive and low-cost diagnosis. In the last few decades, due to the introduction of new and powerful diagnostic tools such as ultrasound and Doppler imaging, as well as due to the high proficiency required for accurate heart auscultation, HS auscultation has been relegated to a minor role in daily medical practice. Conversely, developments in digital signal processing and analysis are leading to renewed interest in HS (an extensive review is presented in Abbas and Bassam (2009) and Durand and Pibarot (1995)). It has emerged as a powerful (easy to use, less intrusive, repeatable and accurate) and inexpensive bio-signal to develop and deploy monitoring systems, mainly in the context of chronic disease management where low-cost and reliable solutions for cardiovascular function assessment are required for long-term patient follow-up. This application scenario has been growing in importance in the past decades due to the rising incidence and prevalence of chronic cardiovascular diseases (CVD) (WHO 2005) as well as the unprecedented ageing of the world population (Rechel *et al* 2009, United Nations 2001) with a decrease in the number of working age per retiree. The solution to this health, social and economical problem is believed to be changing the focus from curative healthcare to preventive long-term healthcare, i.e. controlling costs (social and economical) by reducing preventable healthcare conditions and avoiding/shortening hospitalization. In this sense, long-term tele-monitoring is a promising tool to achieve the aforementioned goal.

HS is a consequence of turbulent blood flow and vibrating cardiovascular structures, which propagate to the chest. These vibrations typically result from myocardial and valvular events that are affected by the hemodynamic and electrical activity of the muscle (Durand and Pibarot 1995). There are two main HSs, S1 and S2, which are related to the systole and the diastole of the heart cycle, respectively. The first HS (S1) occurs at the onset of the ventricular systole. Nowadays, there is still no uniquely accepted theory regarding the origin of this two-component HS (Durand and Pibarot 1995). Its first component is usually attributed to the closing of the atrio-ventricular (AV) valves (mitral and tricuspid), while its second high-frequency component[5] is typically associated with the vibration induced by the left ventricle ejection (Tavel 1967). The second HS (S2) marks the onset of the ventricular diastole of the heart cycle. It is composed of two main high-frequency components, which are highly correlated with the closing of the aortic and the pulmonary valves, respectively.

Existing commercial and research biomedical systems using HS are mainly supported by the analysis of the intensity and spectral content of the main HS components (e.g. Debbal and Bereksi-Reguig 2008, Eitz *et al* 2003, Xiao *et al* 2003, Durand and Pibarot 1995), in tasks such as noise detection (Kumar *et al* 2011) or HS segmentation (Schmidt *et al* 2010,

[5] It is important to notice that both S1 and S2 are, in rigor, low-frequency components (e.g. compared to other sounds such as murmurs). However, comparing S1 and S2, S2 sounds may be regarded as high-frequency components.

Ahlstrom *et al* 2008). Other works combine HS and electrocardiogram (ECG), e.g. Syed *et al* (2007) employ these signals in the detection and analysis of systolic and diastolic intervals. In this paper, the goal is to assess the feasibility to accurately extract the main systolic time intervals. The underlying hypothesis is that the first and the second HSs encode mechanical activity (valve movements, blood flow, etc) and that these components exhibit noticeable and specific signatures that enable their identification using this signal. To this end, we also follow the idea of combining HS and ECG.

Using HS as a reference for systolic and diastolic time-interval measurement is not a new idea. It was common practice prior to the introduction of ultrasound. The procedure was based on the combination of the carotid pulse (variants exist where other markers for the systolic ejection are applied), the HS and the ECG. By neglecting the time distortion caused by pulse propagation, the carotid pulse wave was analyzed to identify the aortic valve timings and enabled the estimation of the left ventricular ejection time (LVET). The pre-ejection period (PEP) was measured indirectly by subtracting LVET from electro-mechanical systole RS2, where RS2 is the time interval from the ECG R-peak to S2. However, one aspect that has not been fully explored using HS is the possibility to accurately measure the main cardiac time intervals using this signal without resorting to the carotid pulse, as is proposed in this paper.

The clinical relevance of this goal stems from the fact that the myocardial relaxation and contraction are governed by intracellular recycling of calcium ions. Therefore, the timings of these basic cardiac functions are directly related to the health of the cardiac cells (Oh and Tajik 2003), which determine the ability to ensure blood delivery according to the metabolic requirements of the organs. Of major importance are the timings of the left ventricle, since it is this ventricle's function to control the blood flow in the systemic circulation. The left ventricle acts as a pump with two main phases: the systolic ejection and the diastolic filling. Clinically accepted descriptions of systolic and diastolic function can be readily obtained using invasive as well as non-invasive procedures (e.g. echocardiography) in clinical settings, which are not adequate for daily applications in home settings as required for long-term patient follow-up. An adequate alternative for this type of application scenario to assess the global cardiac function is the use of time intervals (Oh and Tajik 2003, Finckelstein and Cohn 1993, Werf *et al* 1975, Weissler *et al* 1968). Of major relevance in assessing the left ventricular systolic function are the PEP and the LVET. By definition, PEP is the time interval between the start of ventricular depolarization and the moment of aortic valve opening, whereas the LVET is defined as the time interval of left ventricular ejection, which occurs between the opening of the aortic valve and its subsequent closure.

The remainder of the paper is organized as follows: in section 2 the algorithms for HS analysis are described. The data collection strategy is presented in section 3. In section 4, the main results are presented and discussed. Finally, in section 5 the main conclusions are presented.

2. Methods

2.1. PEP estimation

A Bayesian approach is followed for PEP estimation, resorting to the instantaneous amplitude (IA) of the HS waveform as the main feature. The motivation for this approach comes from the fact that the closure of AV valves is usually correlated with strong amplitude values in the first HS. To bootstrap the method, the IA curve is analyzed and the PEP duration is estimated based on the delay between AV closure and aortic valve opening reported in the literature (Tavel 1967). The PEP interval of the previous heartbeat is also included in the model to constrain the

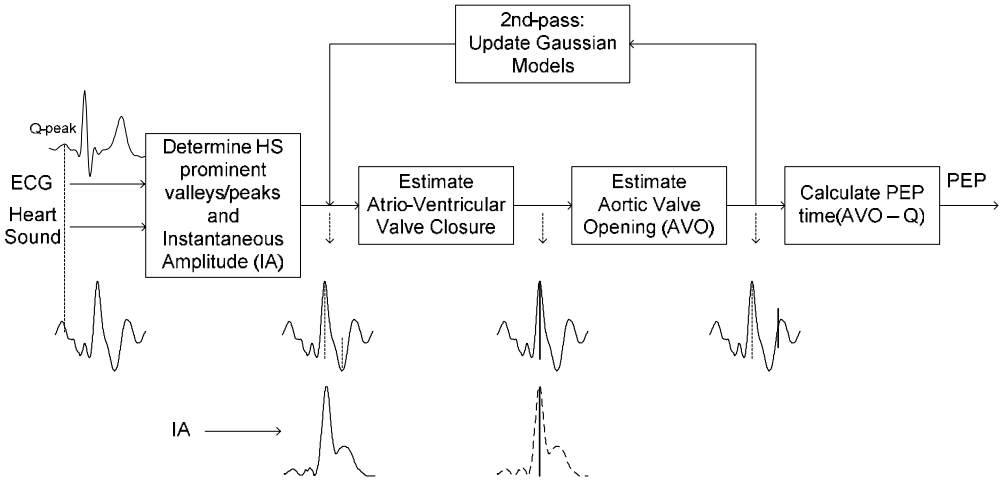

Figure 1. Overview of the proposed PEP estimation approach.

range of possibilities, since it might be assumed that, at rest, abrupt variations are not likely to occur.

Additionally, a two-pass approach for PEP estimation is followed. In the first iteration, initial probability distributions for the variables included in the model are assumed. These are parameterized using average population-based values reported in the literature. Then, the initial distributions are updated based on the results obtained. A second run of the algorithm is then conducted using the updated distributions. Thus, the actual model applied is patient-dependent and is estimated using a data-driven approach.

PEP is formally defined as the time interval between the Q-peak of the ECG to the opening of the aortic valve. Hence, an algorithm for detection of Q-peaks (Henriques *et al* 2008) is applied.

The procedures carried out for PEP estimation are summarized in figure 1. Each of the depicted stages will be described in the next paragraphs.

2.1.1. First pass. Given a HS signal, $s(t)$, the algorithm for PEP estimation starts by determining the signal's IA, $a(t)$, via the analytic signal as in (1). There, $\mathrm{HT}(\cdot)$ denotes the Hilbert transform:

$$a(t) = |s(t) + j\mathrm{HT}\,(s(t))|\,. \tag{1}$$

AV valve closure estimation. Next, for each heartbeat, k, the AV closure time interval, AV_k, is estimated. To this end, the corresponding Q-peak (previously determined) is employed as reference. A Bayesian model is defined, where the prominences, prom_k, of the HS near the Q-peak, the IA curve, IA_k, and the previous AV interval, AV_{k-1}, are employed according to

$$p(\mathrm{AV}_k|\mathrm{prom}_k, \mathrm{IA}_k, \mathrm{AV}_{k-1}) \approx p(\mathrm{AV}_k|\mathrm{prom}_k) \cdot p(\mathrm{AV}_k|\mathrm{IA}_k) \cdot p(\mathrm{AV}_k|\mathrm{AV}_{k-1}). \tag{2}$$

Here, $\mathrm{AV}_k,\ \mathrm{IA}_k,\ \mathrm{AV}_{k-1} \in D_{\mathrm{AV}}$, which denotes the time range for AV time estimation, i.e. the time interval $D_{\mathrm{AV}} = [\text{Q-peak time, Q-peak time} + 110\ \text{ms}]$. This interval was obtained experimentally and proved reliable. In fact, we observed that the AV valves close 44 ± 12 and 51 ± 11 ms after the Q-peak, for the healthy and CVD populations, respectively. These values agree with studies reported in the literature, e.g. von Bibra *et al* (1986). In order to accommodate possible larger variations, the interval was made intentionally broader.

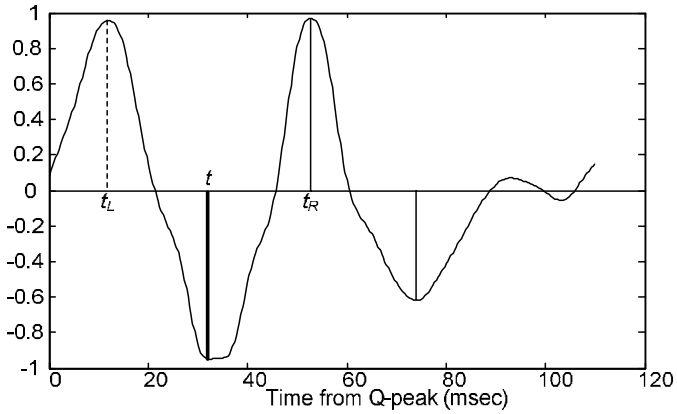

Figure 2. Detection of prominences in the HS (vertical scale is arbitrary).

In (2), in order to model $p(\mathrm{AV}_k \mid \mathrm{prom}_k)$, prominent valleys and peaks of the HS in D_{AV} are first determined, as the closing moment of the AV valves is strongly correlated with those events. Those valleys and peaks are found according to the algorithm described in Paiva *et al* (2008), as in (3). There, t_R and t_L denote the instants of the left and right peaks corresponding to the valley in t (or minima, in the case that the instant t is a peak):

$$\mathrm{prom}_k(t) = \begin{cases} \min\left(|s(t) - s(t_L)|,\ |s(t) - s(t_R)|\right) & t \text{ is peak/valley} \\ 0 & \text{otherwise} \end{cases}, \qquad p(\mathrm{AV}_k|\mathrm{prom}_k) = \frac{\mathrm{prom}_k}{\int_{t\in D_{\mathrm{AV}}} \mathrm{prom}_k \, \mathrm{d}t}. \tag{3}$$

This procedure is illustrated in figure 2. There, the solid curve represents the segment of HS corresponding to D_{AV}, and the vertical solid lines denote found prominences. The left and right peaks corresponding to the prominence in t (bold solid vertical line) are also represented.

Also in (2), the conditional probability distribution of the AV closure time interval given the IA, $p(\mathrm{AV}_k \mid \mathrm{IA}_k)$, is defined as the normalized IA in beat k, according to (4). This is based on the observation that higher amplitude values are more likely to correspond to AV closure (Tavel 1967).

$$p\left(\mathrm{AV}_k|\mathrm{IA}_k\right) = \mathrm{IA}_k \Big/ \int_{t\in D_{\mathrm{AV}}} \mathrm{IA}_k(t). \tag{4}$$

Still in (2), $p(\mathrm{AV}_k \mid \mathrm{AV}_{k-1})$ is modeled as a Gaussian distribution centered in the previous AV interval, AV_{k-1}, and with a standard deviation $\sigma_{\mathrm{AV}} = 30$ ms, as in

$$p\left(\mathrm{AV}_k|\mathrm{AV}_{k-1}\right) = G\left(\mathrm{AV}_{k-1}, \sigma_{\mathrm{AV}}\right). \tag{5}$$

Here, $G(\mu,\sigma)$ denotes a Gaussian function with mean value μ and standard deviation σ. However, in the first heartbeat there is no AV interval reference. Hence, in the first beat a uniform distribution is employed.

The assumption in (5) is motivated by the observation that the AV time interval tends to exhibit invariance with respect to heart rate (Warrington *et al* 1988). Nevertheless, in order to accommodate possible variations, a broad standard deviation was allowed.

Finally, the AV interval is estimated as the time instant that maximizes $p(\mathrm{AV}_k \mid \mathrm{prom}_k, \mathrm{IA}_k, \mathrm{AV}_{k-1})$.

PEP estimation. After AV closure interval estimation, PEP duration, PEP_k, is inferred. Again, a Bayesian strategy is followed, employing the estimated AV interval in beat k, AV_k, and the previous PEP interval in beat $k-1$, PEP_{k-1}, as in

$$p\,(\mathrm{PEP}_k|\mathrm{AV}_k, \mathrm{PEP}_{k-1}) \approx p\,(\mathrm{PEP}_k|\mathrm{AV}_k) \cdot p\,(\mathrm{PEP}_k|\mathrm{PEP}_{k-1})\,. \tag{6}$$

Here, $\mathrm{AV}_{k-1} \in D_{\mathrm{AV}}$ and $\mathrm{PEP}_k, \mathrm{PEP}_{k-1} \in D_{\mathrm{PEP}}$, which denotes the time range PEP time estimation, corresponding to the time interval D_{PEP} = [Q-peak time, Q-peak time + 210 ms]. As before, this interval was obtained experimentally and proved reliable. In fact, we observed average PEP ranges of 77 $\pm$ 9 and 78 $\pm$ 17 ms, for the healthy and CVD populations, respectively. Studies reported in the literature suggest higher average PEP ranges, e.g. 143 $\pm$ 19 ms in von Bibra *et al* (1986). Therefore, in order to accommodate possible larger variations, the interval was made intentionally broader (up to 210 ms). In any case, the algorithm was relatively insensitive to interval variations, as will be presented later.

In (6), the conditional probability distribution of PEP duration given the AV time interval, $p(\mathrm{PEP}_k \mid \mathrm{AV}_k)$, is modeled as a Gaussian centered in $\mathrm{AV}_k + \mu_{\text{PEP-AV}}$ ($\mu_{\text{PEP-AV}} = 30$ ms), again with a standard deviation $\sigma_{\text{PEP-AV}} = 30$ ms, by results found in the literature, indicating that the aortic valve opens typically 30 ms after the closure of AV valves (Tavel 1967):

$$p\,(\mathrm{PEP}_k|\mathrm{AV}_k) = G\,(\mathrm{AV}_k + \mu_{\mathrm{PEP-AV}}, \sigma_{\mathrm{PEP-AV}})\,. \tag{7}$$

As previously explained, this model (as well as the other Gaussian models) only serves for bootstrapping the identification process; later they will be updated using a data-driven approach.

Still in (6), $p(\mathrm{PEP}_k \mid \mathrm{PEP}_{k-1})$ is modeled as a Gaussian distribution centered in the previous PEP interval, PEP_{k-1}, and with a standard deviation $\sigma_{\mathrm{PEP}} = 30$ ms, as in (8). Since there is no PEP interval reference for the first heartbeat, a uniform distribution is initially taken.

$$p\,(\mathrm{PEP}_k|\mathrm{PEP}_{k-1}) = G\,(\mathrm{PEP}_{k-1}, \sigma_{\mathrm{PEP}})\,. \tag{8}$$

The assumption in (8) is motivated by the relative invariance of PEP to heart rate (Warrington *et al* 1988). On the other hand, it is observed that PEP does change with load (Muehlsteff *et al* 2006). At rest no significant changes in load and in heart rate are expected between consecutive heartbeats. Nevertheless, a broad standard deviation was allowed, so as to accommodate possible variations.

Finally, the PEP interval is estimated as the time instant that maximizes $p(\mathrm{PEP}_k \mid \mathrm{AV}_k, \mathrm{PEP}_{k-1})$.

2.1.2. Second pass. After the first iteration of the algorithm, the initial bootstrapping distributions, $p(\mathrm{AV}_k \mid \mathrm{AV}_{k-1})$, $p(\mathrm{PEP}_k \mid \mathrm{PEP}_{k-1})$ and $p(\mathrm{PEP}_k \mid \mathrm{AV}_k)$, are updated using the obtained results. Mean and standard deviation values for each of the employed Gaussian models are, therefore, calculated separately for each patient. Then, the algorithm is run again using the updated models. Moreover, the estimated AV and PEP distributions, $p(\mathrm{AV}_k \mid \mathrm{prom}_k, \mathrm{IA}_k, \mathrm{AV}_{k-1})$ and $p(\mathrm{PEP}_k \mid \mathrm{AV}_k, \mathrm{PEP}_{k-1})$ are added in the second round, resulting in the following equations, (9) and (10). There, p_{fp} denotes the distributions obtained in the first pass of the algorithm. The other distributions are determined as described previously:

$$\begin{aligned} &p(\mathrm{AV}_k|\mathrm{prom}_k, \mathrm{IA}_k, \mathrm{AV}_{k-1}) \\ &\approx p(\mathrm{AV}_k|\mathrm{prom}_k) \cdot\ p(\mathrm{AV}_k|\mathrm{IA}_k) \cdot p_{\mathrm{fp}}(\mathrm{AV}_k|\mathrm{AV}_{k-1}) \cdot\ p_{\mathrm{fp}}(\mathrm{AV}_k|\mathrm{prom}_k, \mathrm{IA}_k, \mathrm{AV}_{k-1}), \end{aligned} \tag{9}$$

$$p(\mathrm{PEP}_k|\mathrm{AV}_k, \mathrm{PEP}_{k-1}) \approx p_{\mathrm{fp}}(\mathrm{PEP}_k|\mathrm{AV}_k) \cdot p_{\mathrm{fp}}(\mathrm{PEP}_k|\mathrm{PEP}_{k-1}) \cdot p_{\mathrm{fp}}(\mathrm{PEP}_k|\mathrm{AV}_k, \mathrm{IA}_k, \mathrm{PEP}_{k-1}). \tag{10}$$

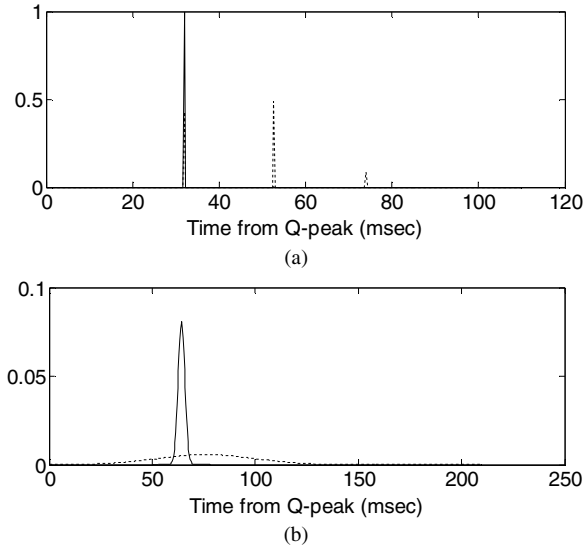

Figure 3. (a) AV closure probability distribution; (b) PEP probability distribution. The dashed lines represent the probability distributions after the first pass of the algorithm. The solid lines denote the final distributions.

An example of the obtained AV closure and PEP probability distributions for one heartbeat are shown in figure 3. There, the dashed lines represent the probability distributions after the first pass of the algorithm, while the solid lines denote the final distributions.

2.2. LVET estimation

As the aortic valve closure corresponds to the start of the S2 HS component, the LVET estimation is based on the detection of the S2 onset.

The approach for LVET estimation is summarized in figure 4.

In the first step of the algorithm, the goal is to segment the HS waveform into a number of sound lobes representing candidate S2 sound segments. This segmentation resorts to the application of the Shannon energy operator to the detail coefficients of the fast wavelet transform (FWT). More precisely, the signal is first high-pass filtered using a 25 Hz cutoff Butterworth filter in order to eliminate low frequency noises (e.g. from muscle movements, etc) and the filtered signal's envelope is then obtained. To calculate the envelope, we compute the FWT of the signal using the db6 (Daubechies wavelet family) because of its ideal shape for capturing the transient nature of the sounds. The FWT is performed until the second level (decomposition depth was experimentally tuned). The obtained approximation coefficients represent the transient part of the low frequency HS and are used to extract the signal envelope based on the following Shannon energy operator (11):

$$E[w] = \frac{1}{N}\sum_{i=1}^{N}(x_w[i])^2 \cdot \log_{10}(x_w[i])^2 \qquad w = 1,\ 2,\ \ldots, W. \tag{11}$$

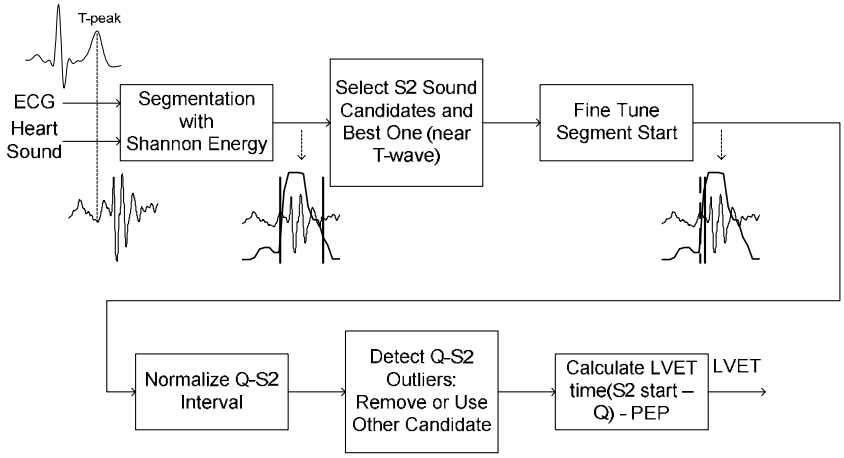

Figure 4. Overview of the proposed LVET estimation approach.

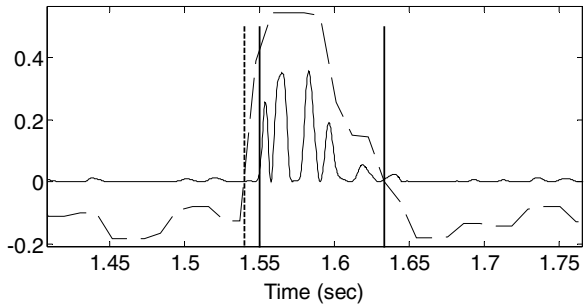

Figure 5. S2 sound segmentation and start adjustment (vertical axis is arbitrarily scaled).

In (11), $x_w[k]$ represents the ith approximation coefficient in a given N-sample window, w. In fact, the Shannon energy is calculated in 20 ms windows with 50% overlap, resulting in a total of W windows. The normalized Shannon energy, E_n, is then computed as follows:

$$E_n[w] = \frac{E[w] - \text{mean}(E)}{\max(E)} \qquad w = 1,\ 2,\ \ldots, W. \tag{12}$$

Finally, sound lobe boundaries are identified from zero-crossings of the normalized Shannon energy. In figure 5 the dashed curve represents the Shannon energy and the vertical lines denote segment boundaries (more details below).

As a result of the above procedure, some segments correspond to S2 sounds, while other spurious segments may correspond to noise. In the case that data from patients with CVD are employed, segments corresponding to murmurs may also appear. This situation is not tackled in this work, but was handled in Kumar *et al* (2006, 2008).

In this paper, a different approach is followed, where we take advantage of the ECG's T-wave. Therefore, after sound segmentation, the goal is to determine the best segments corresponding to the S2 sounds. A robust way to restrict the number of S2 segment candidates in a given heartbeat is to use the T-wave gating, since the S2 sound will start very close to it. As before, the T-peak is obtained using the algorithm introduced in Henriques *et al* (2008). Thus, only the segments starting in the [T-peak − 40 ms, T-peak + 100 ms] interval are considered

as S2 candidates. A broad interval capable of dealing with possible large variations is applied. The obtained S2 segment candidates are then analyzed and the best one is selected according to the following rules.

- The earliest segment is the first S2 candidate, as it is very close to the T-wave peak.
- If there is a later segment with at least double the energy of the first one, then the later one is selected. This procedure is repeated iteratively for the current selected segment until no S2 changes occur.

The original segmentation algorithm (Kumar *et al* 2006) was developed in the context of HS segmentation for application in the detection of prosthetic heart valve dysfunctions. As this task does not entail severe temporal constraints, a somewhat cautious approach was followed for the detection of segment starts and endings. Hence, in this work, the original segmentation algorithm is improved by redefining the start of each S2 sound as the point where the signal's energy reaches 10% of the maximum energy in the segment. Figure 5 illustrates the analysis for one typical S2 sound segment. There, the solid function plot denotes the signal energy, while the dashed function plot represents the normalized Shannon energy. The three vertical lines are, from left to right: initial definition of the S2 sound start, based on the Shannon energy; corrected S2 sound start based on the signal energy; and the end of the S2 sound.

After fine-tuning S2 candidate segment onsets, the Q-S2 intervals are calculated in the sound waveform, normalized to zero heart rate (Warrington *et al* 1988), and possible outliers are identified, which will be replaced by more reliable S2 candidates in terms of the stability of that interval. In fact, in the experimental setup (see section 3), the normalized Q-S2 is expected to be reasonably stable, since patient posture and load do not change noticeably.

The Q-S2 is normalized to zero heart rate according to the equations derived in Warrington *et al* (1988):

$$\mathrm{QS2}_{\mathrm{norm}} = \mathrm{QS2}\,[k] + \mathrm{gain} \cdot \mathrm{HR}\,[k]\,. \tag{13}$$

In (13), HR[k] denotes the instantaneous heart rate at the kth heartbeat, computed based on the R–R interval. QS2[k] and $\mathrm{QS2}_{\mathrm{norm}}[k]$ represent the actual and normalized Q-S2 intervals at beat k. As for the gender-dependent gain, values of 1.2 and 1.3 are employed for males and females, respectively (Warrington *et al* 1988). The $\mathrm{QS2}_{\mathrm{norm}}$ sequence in the sound waveform is then analyzed and outliers are detected. The median value of the sequence, m, is calculated and points k that are at least 5% away from the median are marked as outliers, i.e.

$$\frac{\mathrm{QS2}_{\mathrm{norm}}[k] - m}{m} > 0.05 \Rightarrow k \text{ is an outlier.} \tag{14}$$

Next, each detected outlier is either replaced by another S2 candidate or discarded if no reliable substitute is found, according to (14). We define the best substitute as the segment with the least difference from the normalized Q-S2 interval.

Finally, LVET is obtained by calculating the difference from aortic closing and opening times (the latter obtained from the PEP estimation algorithm).

3. Data collection

3.1. Experimental setup

The data considered in this paper have been obtained using two groups of volunteers: one containing 23 healthy subjects and another with 12 patients suffering from different CVD, such as hypertension, arrhythmia, acute infarction, AV blocks (pacemaker), angina, heart failure, ischemia, aortic insufficiency, coronary artery disease and aortic stenosis. These volunteers

have been asked to participate in the data collection at the Centro Hospitalar de Coimbra (CHC), aimed at the simultaneous acquisition of HS and echocardiography (echo). This data collection study was authorized by CHC's ethical committee and signed consent was collected. A synchronous ECG with each of the above signals was also acquired and served as a reference signal for co-registration.

The population was not balanced for gender. More specifically, in the healthy population, 21 males and only 2 females volunteered. The average HR during data collection was 71.9 $\pm$ 11.9 bpm. The biometric characteristics of the healthy population were

- age: 26.1 $\pm$ 5.6 years,
- BMI: 24.1 $\pm$ 2.6 kg m^{-2}.

As for the CVD population, this one is more balanced for gender, as eight male and four female patients volunteered. The average HR during data collection was 70.1 $\pm$ 11.3 bpm. The biometric characteristics of the CVD population were

- age: 55.7 $\pm$ 18.4 years,
- BMI: 25.6 $\pm$ 3.3 kg m^{-2}.

The measurement protocol was conducted by an authorized medical specialist and consisted of several acquisitions of echocardiography in Doppler mode and HS collected at the left sternum border (LSB). More precisely the following steps were carried out.

- The patient was set in supine position, turned left (approximately 45°)—the usual echo observation position for the aortic valve.
- The echo was configured for Doppler mode and the stethoscope was positioned in the LSB region.
- Runs of 30–60 s. Data acquisitions of HS, echo and ECG were performed repeatedly.

The following signals have been acquired.

- Echocardiography and ECG have been acquired using a Vivid system from General Electric. This device produces outputs with images of 500 Hz time resolution (see figure 6).
- HS and ECG: a Meditron stethoscope and analyzer were applied to record HS and ECG at 44.1 kHz. The bandwidth of the HS sensor is 20 kHz.

3.2. Data synchronization and annotation

As already mentioned, the data streams originating from the HS and the echocardiography were synchronized using the simultaneously acquired ECG signals. The algorithm applied for data registration using both ECG signals is based on the matching of the ECG's R–R intervals under least-squares error minimization. Let $R_1(k)$, $k = 1, \ldots, n$, and $R_2(w)$, $w = 1, \ldots, m$ ($n < m$), be the R–R intervals of the ECGs to be registered. The registration instant t is obtained according to

$$t = \underset{w=1,\ldots,m-n}{\arg\min} \left\{ \frac{1}{n} \sum_{k=1}^{n} (R_1(k) - R_2(k))^2 \right\}. \tag{15}$$

The annotations of the opening and closing instants of the aortic valve were performed using the echocardiographies by an experienced clinical expert. The opening instant of the aortic valve was annotated as the onset of the ejection lobe of the left ventricle, while the closing point was defined immediately before the onset of the closing click produced by the residual reflux after the aortic valve cusps have closed, as can be observed in figure 6.

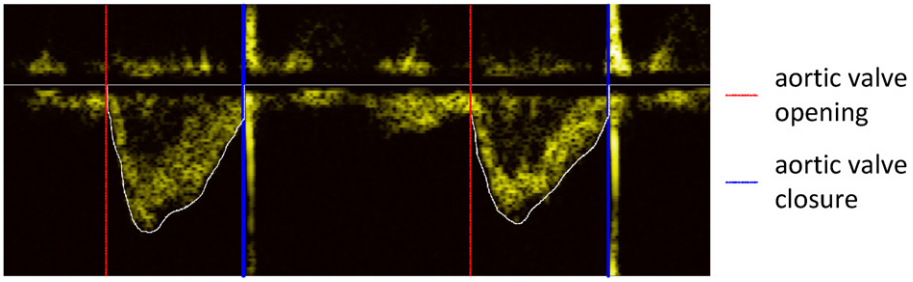

Figure 6. Annotation of aortic valve timings using Doppler mode echocardiography.

Table 1. Summary of results: healthy population.

Parameter	Annotated range (ms) (average ± std)	Estimation error (ms) (average ± std)	ρ
PEP	76.86 ± 9.09	7.67 ± 5.92	0.51
LVET	267.41 ± 25.13	11.39 ± 8.98	0.87

Table 2. Summary of results: CVD population.

Parameter	Annotated range (ms) (average ± std)	Estimation error (ms) (average ± std)	ρ
PEP	77.77 ± 17.61	11.86 ± 8.30	0.68
LVET	292.96 ± 32.68	17.51 ± 17.21	0.83

4. Results and discussion

The main results obtained in this study are summarized in tables 1 and 2, for the healthy and CVD population, respectively. The achieved results suggest that it is possible to accurately identify the systolic time intervals using HS.

4.1. PEP estimation

Regarding the healthy population, 942 annotated heartbeats were acquired. For PEP estimation, 7.67 ms absolute average error, with 5.92 ms standard deviation, resulted, i.e. 9.97% ± 7.7%, relative to the average PEP values (76.86 ms), annotated from the echocardiography. Moreover, 0.51 Pearson's correlation (ρ) between annotated and estimated PEP values was obtained (this was applied as both distributions are Gaussian, from the Kolmogorov–Smirnov test; also, p-values were very low, permitting to discard the null-hypotheses of no correlation). Under the followed protocol, PEP values for subjects in this population are reasonably stable, with slight annotated oscillations between consecutive beats. The output estimated by the proposed algorithm is usually close to the annotated one. However, sometimes the PEP variation is estimated in the opposite direction (possibly due to algorithm or annotation inaccuracies). For this reason, the ρ value is somewhat low.

As for the CVD population, 404 beats were annotated. In terms of PEP estimation, 11.86 ms absolute average error, with 8.30 ms standard deviation, resulted, i.e. 15.25% ± 10.67%, relative to the average annotated PEP values (77.77 ms). In addition, 0.68 correlation was attained.

Comparing to the healthy population, a higher estimation error is observed. This is mostly a consequence of a more complex sound signal morphology in this population, resulting from

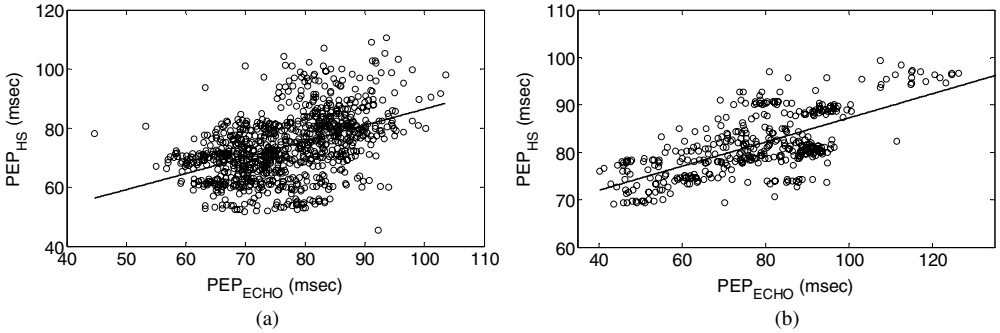

Figure 7. Measurement sensitivity for PEP. (a) Healthy population ($\rho = 0.51$): best linear fit $y = 0.55x + 31.89$. (b) CVD population ($\rho = 0.68$): best linear fit $y = 0.25x + 61.78$.

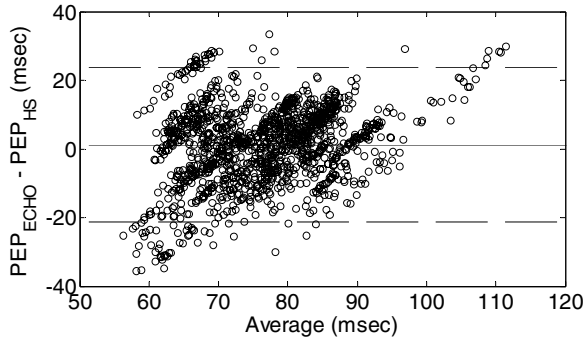

Figure 8. Bland–Altman scatter plot of the estimated beat-to-beat PEP estimation with respect to echocardiography annotation $(\mathrm{PEP_{ECHO}} + \mathrm{PEP_{HS}})/2$.

higher average BMI, age and blood flow issues related to the patient condition. For instance, body fat acts like a low-pass filter as well as a gain attenuator for HS.

In any case, the estimation error is still low, especially compared to other approaches, as will be discussed later. On the other hand, the ρ value is higher. The abovementioned PEP oscillations are still observed, but since PEP variability is higher in this population, estimation variations tend to follow the same direction as the annotated ones.

Figure 7 shows PEP measurement sensitivity. For the healthy population (figure 7(a)), the best linear fit corresponds to a 0.55 slope, which is far from unity, the theoretical ideal. This seems to result from the noticeable dispersion of estimated PEP values for any given annotation value, as observed in figure 7(a). Figure 7(b) presents results for the CVD group. There, the best linear fit corresponds to a 0.25 slope, also far from unity. This stems from increased average estimation error for low and upper ends of the PEP range.

The PEP estimation difference dispersion as a function of the beat-to-beat values from echocardiography and HS is depicted in figure 8 (Bland–Altman plot). The horizontal solid line represents mean and the dashed lines denote the level of agreement (mean difference $\pm$ 2 standard deviation). As can be observed, the average error is close to zero (more precisely, 1.19 ms), with 11.28 ms standard deviation. Moreover, the estimation error distributes evenly for a high range of PEP durations.

Table 3. PEP estimation results with different number of iterations.

Parameter	Estimation error (ms) (average ± std)	ρ
PEP 1-pass	8.665 ± 6.641	0.457
PEP 2-pass	7.666 ± 5.916	0.512
PEP 3-pass	7.603 ± 5.884	0.514

To evaluate the effect of the two-pass PEP estimation approach, the results for the single-pass case were computed using the healthy population. Additionally, the outcome of conducting three iterations is also assessed. Summary results are presented in table 3, which prove that PEP estimation accuracy indeed increases noticeably (11.5%) when subject-dependent Gaussian models are obtained and applied in the second pass. The addition of a third iteration produces nearly the same results with increased computational cost. Hence, this possibility is discarded. Similar conclusions are obtained for the CVD population.

In order to assess the sensitivity of the algorithm to parameter variations, the results using different initial Gaussian models and only one pass of the algorithm were evaluated. Therefore, the mean difference between AV closure and aortic valve opening was varied up to ± 15 ms from the nominal value. Also, the standard deviations of all Gaussians were varied in the same range. As for the standard deviations, these had nearly null impact on the results: the maximum observed average error was 9 ms. Regarding variations of the mean, these had a more significant impact on the results as expected: a 45 ms mean average value led to 14.1 ms error. Thus, the achieved results seem to confirm Tavel's indication that the aortic valve opens typically 30 ms after the closure of AV valves. However, it should be pointed out that this could change with, e.g., load variations. Hence, a broad standard deviation was employed in the bootstrapping Gaussian distribution to accommodate larger variations from the average.

In the proposed method, only one feature (IA) was utilized. Other dynamics involved in the aortic valve opening process could probably be captured by other features, namely the instantaneous frequency (IF). In order to evaluate this hypothesis, we incorporated the IF into the model, estimated both resorting to the analytic signal and the Wigner–Ville distribution (WVD) (Auger *et al* 1996). However, the results were nearly the same or slightly worse. In fact, IF estimation based on the analytic signal can only be accurate when there is only a single frequency present at each time instant, which is not the case in HS. On the other hand, the WVD allows a time–frequency representation of multi-component non-stationary signals without the time–frequency resolution trade-off limitation of the short-time Fourier transform. However, many interference cross-terms appear in the time–frequency representation, which poses serious difficulties in the accurate estimation of the actual frequencies present at each time instant. A few WVD extensions that dispose off those cross-terms were evaluated, e.g. the pseudo-WVD. However, this is attained at the expense of diminished frequency resolution. All in all, no benefits to PEP estimation resulted, with the drawback of a heavier model. Therefore, this hypothesis was abandoned.

Besides the difficulty in accurate IF estimation, there are other possible causes for the obtained PEP estimation error. Namely, the signal-to-noise ratio (SNR) in several audio waves was lower than the ideal. In any case, the PEP results are quite promising, especially when compared with other state-of-the art methodologies such as impedance cardiography (ICG). In fact, in a comparative study carried out by the team (Carvalho *et al* 2010), PEP estimation error using the proposed HS-based algorithm was 27.4% lower than the best performing ICG-based approach. It is important to mention here that the study was conducted using a young and healthy group of subjects. Elderly populations tend to exhibit much stiffer arteries and higher

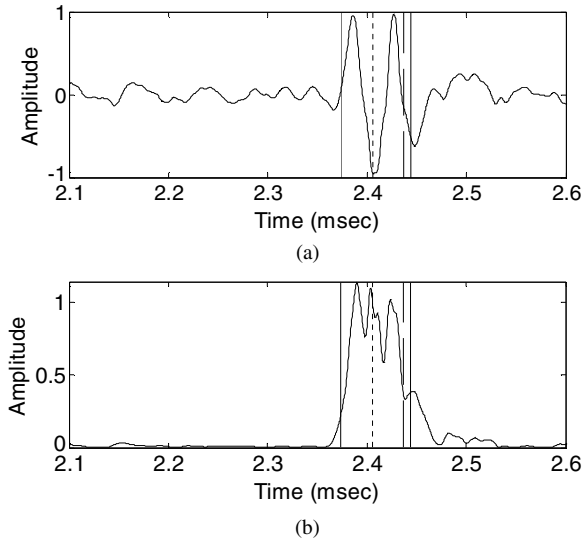

Figure 9. Illustration of PEP estimation. (a) HS waveform. (b) IA. The vertical lines (from left to right): R-peak, AV closure, estimated and annotated aortic valve opening.

blood pressures, so it is likely that the performance of ICG-based approaches will decrease in this population.

Although our results suggest that HS might lead to better systolic time interval (STI) estimation accuracy, the estimated average PEP error could be a clinical issue. In fact, the achieved 15.25% $\pm$ 10.67% average estimation error may lead to inaccurate estimation of cardiac function parameters such as stroke volume or contractility. We conducted a study on the estimation of such parameters based on the systolic time intervals estimated from HS (Couceiro *et al* 2011). In terms of stroke volume, preliminary results indicate 10% $\pm$ 9% estimation error, a value substantially below the clinically accepted error of 30% (Critchley and Critchley 1999).

Figure 9 illustrates typical results of PEP estimation for one heartbeat. The vertical lines denote (from left to right): R-peak, AV closure, estimated and annotated aortic valve opening. As can be observed, the closure of AV valves is accurately detected (a prominent valley in the sound waveform matching a strong peak in the IA curve—figure 9(b)) and the estimated aortic valve opening is close to the one annotated via the echocardiography (around 7 ms).

4.2. LVET estimation

As for LVET estimation from HS, 718 beats from the healthy population have been utilized (a few less beats compared to PEP due to annotation difficulties of aortic valve closure). As can be seen in table 1, 11.39 ms average absolute error was achieved, with 8.98 ms standard deviation, i.e. 4.26% $\pm$ 3.36%, relative to the average annotated LVET values (267.41 ms). Pearson's correlation between annotated and estimated LVET values was 0.87.

As for the CVD population, 174 beats were annotated. From table 2, the proposed LVET estimation algorithm attained 17.51 ms absolute average error, with 17.21 ms standard deviation, i.e. 5.98% $\pm$ 5.87%, relative to the average annotated LVET values (292.96 ms). Furthermore, 0.83 Pearson's correlation was attained.

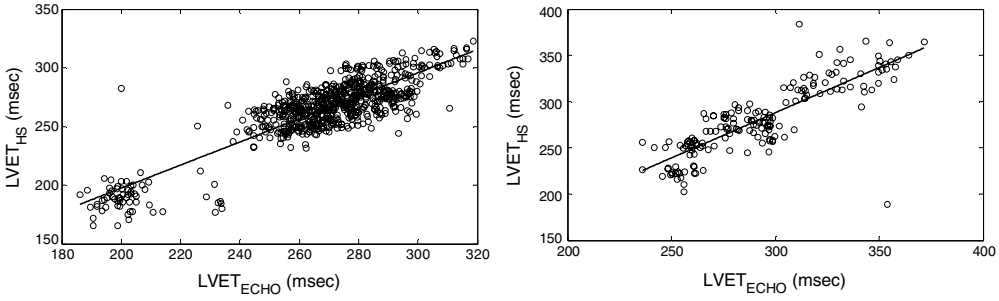

Figure 10. Measurement sensitivity for LVET. (a) Healthy population ($\rho = 0.87$): best linear fit $y = 0.98x + 0.65$. (b) CVD population ($\rho = 0.83$): best linear fit $y = 0.98x - 5.91$.

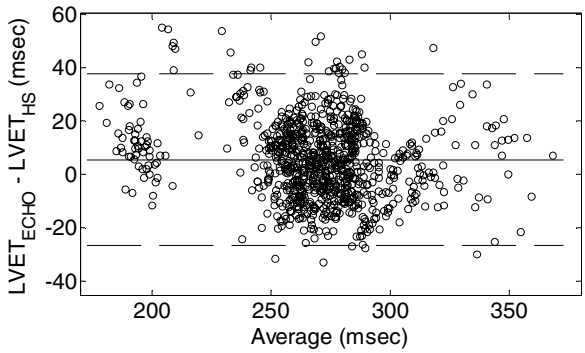

Figure 11. Bland–Altman scatter plot of the estimated beat-to-beat LVET estimation with respect to echocardiography annotation $(\mathrm{LVET_{ECHO}} + \mathrm{LVET_{HS}})/2$.

Comparing the two populations, a higher estimation error is observed in the CVD group. This is a direct consequence of PEP error propagation to LVET estimation. In fact, comparing only the accuracies in aortic valve closure estimation, the algorithm achieves 9.88 and 11.74 ms absolute average error for the healthy and CVD groups, respectively. In any case, both LVET error ranges are similar, in percentage terms.

Comparing to aortic valve opening, aortic valve closure detection is a less difficult problem, tackled by the detection of S2 sound start. As this event can be determined with less uncertainty, correlation values are significantly higher compared to PEP, namely 0.91 in both groups. After subtracting PEP, the attained LVET correlation drops slightly to 0.87 in the healthy group and a bit more (0.83) in the CVD population.

Figure 10 shows LVET measurement sensitivity. For the healthy group (figure 10(a)), the best linear fit corresponds to a 0.98 slope, which is very close to unity. This agrees with the lower LVET estimation errors, as well as higher correlation values. Figure 10(b) depicts results for the CVD population. The best linear fit corresponds to a 0.98 slope, the same as the one obtained in the healthy group. Again, this goes in accordance with the lower LVET estimation errors, as well as higher correlation values.

The LVET estimation difference dispersion as a function of the beat-by-beat values from the echocardiography and HS is shown in figure 11. As previously, the horizontal solid line represents mean and the dashed lines denote the level of agreement. As can be observed, the

average error is close to zero (more precisely, 5.42 ms), with 16 ms standard deviation. In addition, the estimation error is distributed evenly for a high range of LVET durations.

The obtained LVET estimation errors stem from several sources. Besides the mentioned difficulties with the SNR in some acquisitions, the annotated aortic valve closure is persistently delayed compared to the start of sound S2. In a previous study by the team (Carvalho *et al* 2009), it was observed that on average the onset of the aortic valve closing movement was detected 12.1 ms earlier compared to echocardiography. This can be attributed to the fact that HS enables the detection of the onset of the aortic valve closing process, while Doppler echocardiography enables its detection near the closing click induced by the valve cusps, i.e. at the end of the dynamic process. This suggests that the onset of the S2 sound might actually correspond to events associated with blood deceleration rather than aortic cusp movements. Also, LVET estimation error suffers from propagation of PEP detection errors. As a matter of fact, calculating LVET using the annotated PEP values led to the decrease of the average absolute error to 9.88 $\pm$ 8.65 and 11.74 $\pm$ 15.33 ms, in the healthy and CVD groups, respectively. Naturally, the drop is clearer in the CVD population. Taking into consideration the observed average annotation delay, such errors would decrease even more to 7.8 and 11.2 ms, respectively.

5. Conclusions

This paper investigates the possibility of using HS to accurately measure the main systolic heart time intervals, i.e. the PEP and the LVET. The working hypothesis is that HS encodes markers that enable the detection of the opening and closing of the aortic valve. To evaluate this hypothesis a comparative echocardiography–HS study was conducted on 23 healthy and 12 CVD subjects. An automated HS annotation algorithm for the detection of the aortic valve events was described. PEP was estimated following a Bayesian approach where the IA of the HS and the typical delay between aortic valve opening and AV valve closure were employed as the main features. Regarding LVET, sound segmentation was performed (based on the application of the Shannon energy operator to the detail coefficients of the FWT) and segments near the peak of T-wave are taken as S2 sound candidates.

The obtained results strongly support the view that HS can be applied to detect the onset of the aortic valve movement processes. This seems to be a significant achievement since other competing approaches for LVET and PEP measurement (e.g. the ICG approach) tend to exhibit biases in the estimation of those moments, leading to possible inaccuracies in cardiac function assessment. In fact, as already mentioned, there is ample evidence that ICG does not enable a precise detection of the onset of the aortic valve opening and closing process (Ermishkin *et al* 2007, Carvalho *et al* 2010).

The main current limitation of the proposed method pertains to PEP estimation as the opening of the aortic valve is more difficult to detect than its closure. Nevertheless, a recent study (Couceiro *et al* 2011) suggests that cardiac parameters, namely stroke volume, estimated based on the STIs obtained from the present method, provide sufficient clinical accuracy.

In the future, we plan to perform hemodynamic assessment for several distinct CVD, studying the impact of using HS and other competing approaches, namely the ICG-based and PPG-based methodologies, on several application scenarios (hospital, health, etc).

Acknowledgments

This work was supported by the European Integrated Project HeartCycle (FP7-216695) and SoundForLife (PTDC/EEA-ACR/68887/2006). The authors want to express their gratitude

to the volunteers who participated in this study. They would also like to recognize and to express their appreciation to the Centro Hospitalar de Coimbra for supporting the study.

References

Abbas A and Bassam R 2009 *Phonocardiography Signal Processing* (San Rafael, CA: Morgan and Claypool Publishers)

Ahlstrom C, Länne T, Ask P and Johansson A 2008 A method for accurate localization of the first heart sound and possible applications *Physiol. Meas.* **29** 417–28

Auger F, Flandrin P, Gonçalves P and Lemoine O 1996 Time–frequency toolbox—for use with Matlab (France: CNRS; USA: Rice University)

Carvalho P, Paiva R P, Couceiro R, Henriques J, Antunes M, Quintal I and Muehlsteff J 2010 Comparison of systolic time interval measurement modalities for portable devices *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 606–9

Carvalho P, Paiva R P, Couceiro R, Henriques J, Quintal I, Muehlsteff J, Aubert X and Antunes M 2009 Assessing systolic time-intervals from heart sound: a feasibility study *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 3124–8

Couceiro R, Carvalho P, Paiva R P, Henriques J, Antunes M, Quintal I and Muehlsteff J 2011 Beat-to-beat cardiac output inference using heart sounds *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 5657–61

Critchley L and Critchley J 1999 A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques *J. Clin. Monit. Comput.* **15** 85–91

Debbal S and Bereksi-Reguig F 2008 Computerized heart sounds analysis *Comput. Biol. Med.* **38** 263–80

Durand L-G and Pibarot P 1995 Digital signal processing of the phonocardiogram: review of the most recent advancements *Crit. Rev. Biomed. Eng.* **23** 163–219

Eitz T, Fritzsche D, Grimmig O, Frerichs I, Frerichs A, Hellige G, Minami K and Körfer R 2003 Acoustic phenomena and valve dysfunction in cardiac prostheses: data acquisition and collection via the internet *J. Heart Valve Dis.* **12** 414–9

Ermishkin V, Lukoshkova E, Bersenev E, Saidova M, Shitov V, Vinogradova O and Khayutin V 2007 Beat-by-beat changes in pre-ejection period during functional tests evaluated by impedance aortography: a step to a left ventricular contractility monitoring *Proc. 13th Int. Conf. on Electrical Bioimpedance and 8th Conf. on Electrical Impedance Tomography* pp 655–8

Finckelstein S and Cohn J 1993 Method and apparatus for measuring cardiac output *US Patent* 5241966

Henriques J, Carvalho P, Harris M, Antunes M, Couceiro R, Brito M and Schmidt R 2008 Assessment of arrhythmias for heart failure management *Proc. phealth2008—International Workshop on Wearable Micro and Nanosystems for Personalised Health*

Kumar D, Carvalho P, Antunes M and Henriques J 2008 Heart murmur recognition and segmentation by complexity signatures *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 2128–32

Kumar D, Carvalho P, Antunes M, Henriques J, Eugénio L, Schmidt R and Habetha J 2006 Detection of S1 and S2 heart sounds by high frequency signatures *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 1410–6

Kumar D, Carvalho P, Antunes M, Paiva R P and Henriques J 2011 Noise detection during heart sound recording using periodicity signatures *Physiol. Meas.* **32** 599–618

Muehlsteff J, Aubert X and Schuett M 2006 Cuffless estimation of systolic blood pressure for short effort bicycle tests: the prominent role of the pre-ejection period *Proc. Int. Conf. of the IEEE Engineering in Medicine and Biology Society* pp 5088–92

Oh J and Tajik J 2003 The return of cardiac time intervals: the Phoenix is rising *J. Am. Coll. Cardiol.* **42** 1471–4

Paiva R P, Mendes T and Cardoso A 2008 From pitches to notes: creation and segmentation of pitch tracks for melody detection in polyphonic audio *J. N. Music Res.* **37** 185–205

Rechel B, Doyle Y, Grundy E and McKee M 2009 *How Can Health Systems Respond to Population Ageing?* (World Health Organization)

Schmidt S E, Holst-Hansen C, Graff C, Toft E and Struijk J J 2010 Segmentation of heart sound recordings by a duration-dependent hidden Markov model *Physiol. Meas.* **31** 513–29

Syed Z, Leeds D, Curtis D, Nesta F, Levine R and Guttag J 2007 A framework for the analysis of acoustical cardiac signals *IEEE Trans. Biomed. Eng.* **54** 651–62

Tavel M 1967 *Clinical Phonocardiography and External Pulse Recording* (Chicago, IL: Year Book Medical Publishers)

United Nations 2001 *World Population Aging: 1950–2050* (New York: United Nations Publications)

von Bibra H, Wirtzfeld A, Hall R, Ulm K and Blömer H 1986 Mitral valve closure and left ventricular filling time in patients with VDD pacemakers—assessment of the onset of left ventricular systole and the end of diastole *Br. Heart J.* **55** 355–63

Warrington S J, Weerasuriya K and Burgess C D 1988 Correction of systolic time intervals for heart rate: a comparison of individual with population derived regression equations *Br. J. Clin. Pharmacol.* **26** 155–65

Weissler A, Harris W and Schoenfeld C 1968 Systolic time intervals in heart failure in man *Circulation* **37** 149–59

Werf F, Piessens J and Kesteloot H 1975 A comparison of systolic time intervals derived from the central aortic pressure and from the external carotid pulse tracing *Circulation* **51** 310–6

WHO 2005 *Preventing Chronic Diseases: a Vital Investment* (Geneva: World Health Organization)

Xiao S, Guo X, Wang F, Xiao Z, Liu G, Zhan Z and Sun X 2003 Evaluating two new indicators of cardiac reserve *IEEE Eng. Med. Biol. Mag.* **4** 147–52