SUPPLEMENTAL MATERIAL

Data S1.

MARKOV MODEL TO DETECT ATRIAL FIBRILLATION FROM PULSE INTERALS

This section describes the details of the Markov model used to determine the likelihood of AF from a sequence of inter-pulse-intervals (IPI). The IPI sequence is modeled as an 11-states first-order Markov process in which each observed IPI is associated to one of 11 different possible states. This association is done firstly by determining the ratio between the observed IPI and the average IPI in the sequence, and secondly by classifying such ratio into 11 categories of values ranging from 0.75 to 1.25 at steps of 0.05. This corresponds to identify to which extent the observed IPI is shorter, similar, or larger than average. In accordance with the work of Moody and Mark[1], the average IPI ($\overline{IPI}$) was determined recursively as follows:

$$\overline{IPI}(i) = 0.75 \times \overline{IPI}(i-1) + 0.25 \times IPI(i)$$

ECG recordings were used in a previous study[2] as training data to define the statistics of transition from one state to another and derive transition probability matrixes in either presence or in absence of AF. Assuming a sequence **X** of *n* states:

$$\boldsymbol{X} = \{x_1, x_2, x_3, \dots x_n\} \quad (x_i in \{s_{75}, s_{80}, s_{85}, \dots s_{125}\})$$

where *s* indicates one of the 11-states of the model, the transition probability matrix in presence of AF or absence of AF (*not_AF*) can be defined as:

$$p_{j,k,AF} = P\left(x_j \mid x_k, AF\right)$$

$$p_{j,k,not\_AF} = P(x_j \mid x_k, not\_AF)$$

Where x_j and x_k are two consecutive states in the process and $p_{j,k,R}$ their relative frequency during a certain heart rhythm *R* (*AF* or *not_AF*). From this we can derive the conditional probability of **X** given the process *R* (either *AF* or *not_AF*):

$$P(X \mid R) = \prod_{i=1}^{n-1} p_{i+1,i,R}$$

The rhythm *R* that maximizes *P* indicates the maximum likelihood procedure for the observed sequence of states **X**. As suggested by Moody and Mark[1], the logarithm of the transition matrix for not_AF was divided by that of AF and used to determine the Markov score (S_i) iteratively by adding each element in the resulting matrix according to the transitions between the *n* observed states **X**.

$$S_i = \sum_{i=1}^{n-1} \log\left(\frac{p_{i+1,i,not\_AF}}{p_{i+1,i,AF}}\right)$$

In our work *n* was kept equal to the entire length of the IPI sequence. As a final step the Markov score was processed using a previously published averaging filter[1]. The resulting Markov score (S_i) tended to decrease below zero in presence of AF and such value was compared to a threshold to make a binary decision between presence and absence of AF. The Markov model which was designed and trained using ECG data was optimized in our study to process IPI sequences from PPG for AF detection. Such optimization focused solely on fine-tuning the probability threshold applied to the Markov score (S_i) to identify episodes of AF.

Supplemental References:

1. Moody GB, Mark RG. A New Method for Detecting Atrial Fibrillation Using RR Intervals. *Computers in Cardiology*. 1983. p. 227–230.

2. Babaeizadeh S, Gregg RE, Helfenbein ED, Lindauer JM, Zhou SH. Improvements in atrial fibrillation detection for real-time monitoring. *J Electrocardiol.* 2009;42:522–526.