

# UCLA

## UCLA Previously Published Works

### Title

Algorithm for Reliable Detection of Pulse Onsets in Cerebral Blood Flow Velocity Signals.

### Permalink

<https://escholarship.org/uc/item/8m8705jn>

### Authors

Canac, Nicolas

Ranjbaran, Mina

OBrien, Michael

et al.

### Publication Date

2019

### DOI

10.3389/fneur.2019.01072

Peer reviewed



# Algorithm for Reliable Detection of Pulse Onsets in Cerebral Blood Flow Velocity Signals

Nicolas Canac<sup>1\*</sup>, Mina Ranjbaran<sup>1</sup>, Michael J. O'Brien<sup>1</sup>, Shadnaz Asgari<sup>2</sup>, Fabien Scalzo<sup>3</sup>, Samuel G. Thorpe<sup>1</sup>, Kian Jalaleddini<sup>1</sup>, Corey M. Thibeault<sup>1</sup>, Seth J. Wilk<sup>1</sup> and Robert B. Hamilton<sup>1</sup>

<sup>1</sup> Neural Analytics, Inc., Los Angeles, CA, United States, <sup>2</sup> Biomedical Engineering Department and Computer Engineering and Computer Science Department, California State University, Long Beach, CA, United States, <sup>3</sup> Department of Neurology and Computer Science, University of California, Los Angeles, Los Angeles, CA, United States

## OPEN ACCESS

### Edited by:

Bryan G. Young,  
London Health Sciences Centre,  
Canada

### Reviewed by:

Benjamin Aaron Emanuel,  
University of Southern California,  
United States  
Minjee Kim,  
Northwestern University,  
United States

### \*Correspondence:

Nicolas Canac  
nicolas@neuralanalytics.com

### Specialty section:

This article was submitted to  
Neurocritical and Neurohospitalist  
Care,  
a section of the journal  
Frontiers in Neurology

**Received:** 08 April 2019

**Accepted:** 23 September 2019

**Published:** 11 October 2019

### Citation:

Canac N, Ranjbaran M, O'Brien MJ, Asgari S, Scalzo F, Thorpe SG, Jalaleddini K, Thibeault CM, Wilk SJ and Hamilton RB (2019) Algorithm for Reliable Detection of Pulse Onsets in Cerebral Blood Flow Velocity Signals. *Front. Neurol.* 10:1072. doi: 10.3389/fneur.2019.01072

Transcranial Doppler (TCD) ultrasound has been demonstrated to be a valuable tool for assessing cerebral hemodynamics via measurement of cerebral blood flow velocity (CBFV), with a number of established clinical indications. However, CBFV waveform analysis depends on reliable pulse onset detection, an inherently difficult task for CBFV signals acquired via TCD. We study the application of a new algorithm for CBFV pulse segmentation, which locates pulse onsets in a sequential manner using a moving difference filter and adaptive thresholding. The test data set used in this study consists of 92,012 annotated CBFV pulses, whose quality is representative of real world data. On this test set, the algorithm achieves a true positive rate of 99.998% (2 false negatives), positive predictive value of 99.998% (2 false positives), and mean temporal offset error of  $6.10 \pm 4.75$  ms. We do note that in this context, the way in which true positives, false positives, and false negatives are defined carries some nuance, so care should be taken when drawing comparisons to other algorithms. Additionally, we find that 97.8% and 99.5% of onsets are detected within 10 and 30 ms, respectively, of the true onsets. The algorithm's performance in spite of the large degree of variation in signal quality and waveform morphology present in the test data suggests that it may serve as a valuable tool for the accurate and reliable identification of CBFV pulse onsets in neurocritical care settings.

**Keywords:** transcranial doppler, ultrasound, cerebral blood flow velocity, pulse heuristic algorithms, biomedical signal processing

## 1. INTRODUCTION

Assessment of cerebrovascular function is imperative in the diagnosis and management of numerous conditions common in neurological care. Transcranial Doppler (TCD) ultrasound has been used since the 1980s to measure cerebral blood flow velocity (CBFV), which can serve as a means of assessing cerebral hemodynamics (1, 2). TCD has established clinical indications for subarachnoid hemorrhage, cerebral vasospasm, acute ischemic stroke, stenosed or occluded intracranial vessels, and sickle cell disease (3–5). In addition, monitoring of CBFV may be useful as a tool for noninvasive intracranial pressure monitoring and the assessment of mild traumatic brain injury (6–10).

CBFV waveform analysis often utilizes individual CBFV pulse morphology and thus depends heavily on reliable pulse onset detection. Accurate pulse delineation in CBFV waveforms presents a significant challenge for a number of reasons. There is an inherent difficulty in TCD measurement due to the scatter and attenuation of the signal by the skull, resulting in a relatively low signal to noise ratio (11, 12). Additionally, TCD is highly operator dependent and relies on the operator's ability to locate the acoustic window and insonate the appropriate vessel within a cerebrovasculature which may vary widely from patient to patient. Furthermore, CBFV signals are particularly prone to noise artifacts as a result of motion of the (not uncommonly handheld) probe or subject. Finally, the wide range of possible waveform morphologies that can occur in CBFV signals, particularly in pathological populations, presents yet another challenge, as it may be difficult to rely on the presence of any one uniform and consistent feature to aid in onset detection.

Despite these difficulties, TCD remains a compelling diagnostic tool due to being relatively inexpensive, noninvasive, fast, and portable. A solution to the problem of extracting individual CBFV pulses coupled with a reliable and accessible method of CBFV monitoring could open up numerous opportunities for improved neurocritical care. In this study, we focus on the former problem of accurate beat-to-beat segmentation. The problem can be reduced to locating the starting point of each beat or pulse, the so-called pulse onset, as we will refer to it, or pulse foot (13).

Though a number of methods have been developed for pulse onset detection for other physiological signals, including arterial blood pressure (ABP) and photoplethysmogram (PPG) (14–18), relatively few methods have been applied to CBFV signals. Those that do often require information from another complementary signal, such as electrocardiogram (ECG) (19). Several promising methods of CBFV onset detection that require no other complementary signals have been compared, with reported true positive rate (TPR) and positive predictive values (PPV) that can range between around 65% and 98% for both metrics, depending on the choice of error threshold and algorithm parameter selection Asgari et al. (20, 21).

In this paper, we present a new algorithm for pulse onset detection in CBFV waveforms that does not require any additional complementary signals and seeks to improve on the performance of previous methods. This algorithm incorporates a moving difference filter (MDF) and adaptive thresholding to identify windows in which onsets are likely to occur. The algorithm then performs a local search within each window to determine the most probable position for the onset. Finally, the algorithm seeks to identify and correct errors by applying outlier detection strategies. The result of this method is a substantial improvement over existing pulse onset detection algorithms.

We begin by describing the data set that was used to test the performance of the algorithm. We then present a detailed description of the process of defining search windows, locating likely pulse onsets within those windows, and handling onset detection errors. We then move on to a presentation of the performance of the algorithm on the test data set, including an analysis of the impact of free parameters. Finally, we conclude

by summarizing these results and offer some commentary on the potential use cases and robustness of the algorithm.

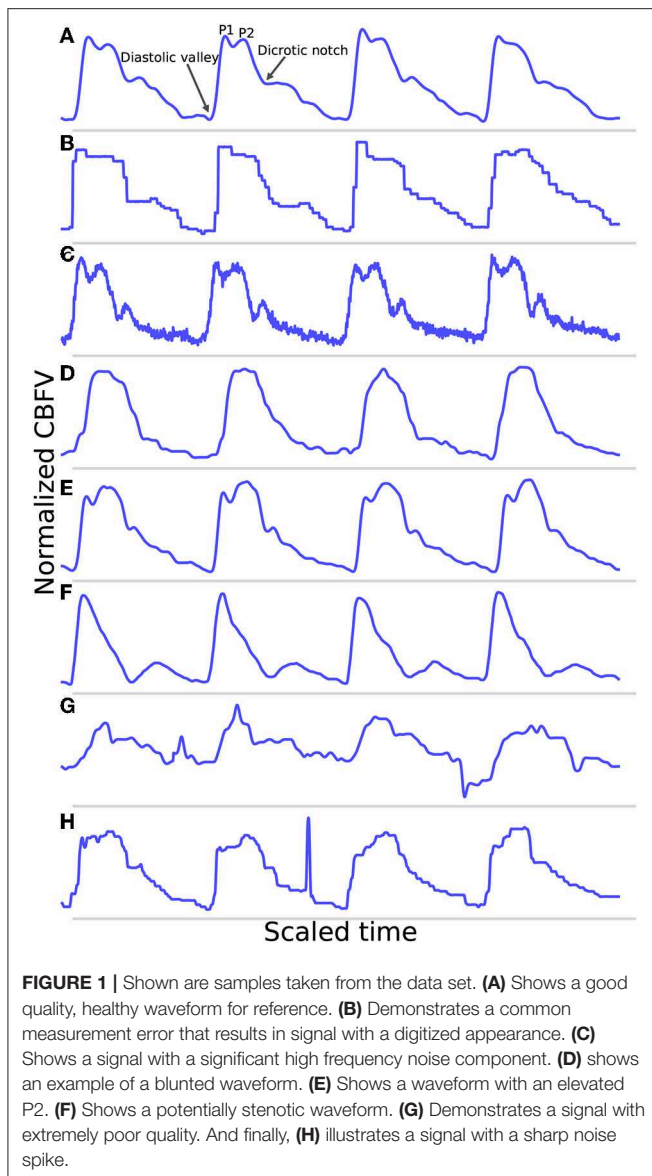
## 2. DATA SET

The data set used in this work is identical to the one used in Asgari et al. (20, 21) and consists of 92,794 annotated CBFV pulses collected by a trained sonographer from subarachnoid hemorrhage patients admitted to the UCLA Medical Center. The subject group consisted of 108 patients (42 female), aged 30–64, with an average age of 48. However, due to unusable or unrecognizable CBFV signals resulting from extreme aliasing and other user errors, six total scans were excluded from this analysis, resulting in a total of 92,012 annotated pulses. Electrocardiogram (ECG) signals were also collected, and, being a simpler waveform for pulse detection, were used for the purpose of facilitating the process of annotating pulse onsets. All patients were consented under the protocol approved by the UCLA Internal Review Board (#10-001331) prior to analyzing any of their data.

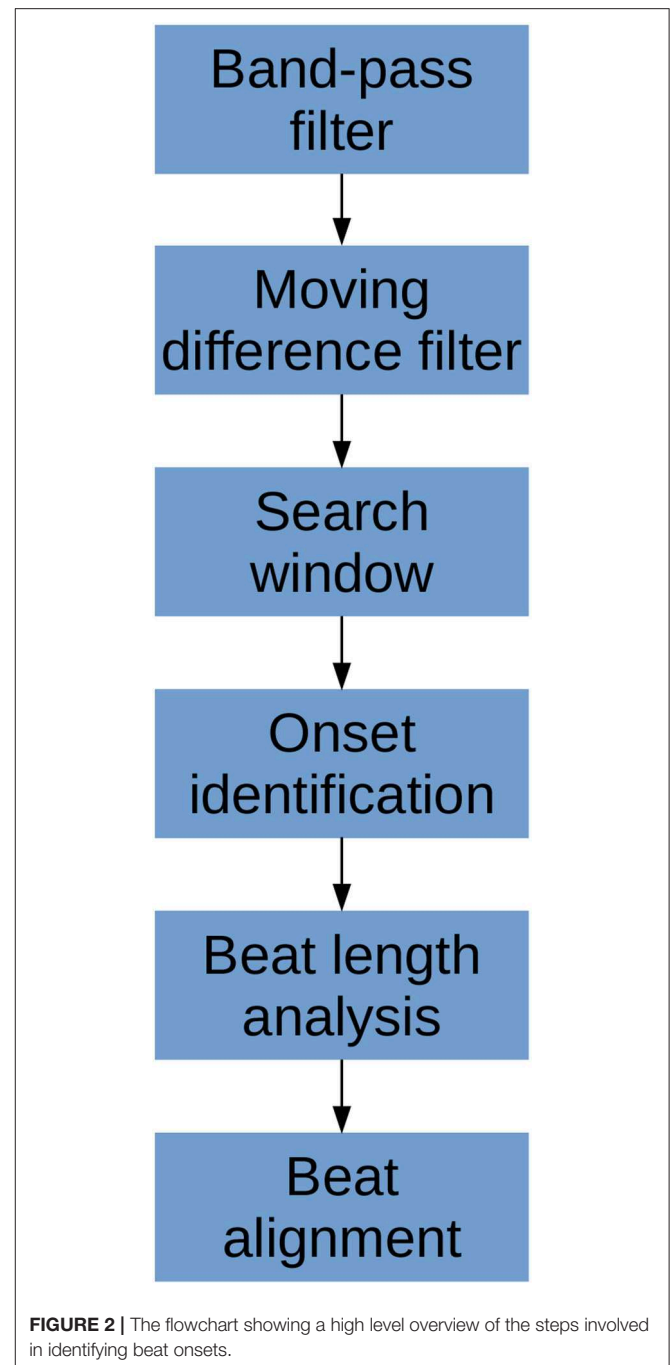
The ground truth for onset locations was determined using the following process. Onset locations were first found using the R peak of the QRS complex in the ECG signal as discussed in Kazanavicius et al. (13) and Asgari et al. (20). The resulting annotations were then manually inspected for accuracy and corrected where necessary. These annotations are treated as the ground truth when evaluating the performance of the algorithm. In order to guard against systematic bias in the onset annotation process, the data set was split in half and a different, independent annotator was assigned to each half. This combination of using a complementary ECG signal along with two independent annotators should provide protection against possible bias within the annotated onsets, and any bias that is present should be detectable as an observable systematic difference between the two halves of the data, none of which was observed. We note that the ECG signal was only used to assist in verification of onset locations, and is not required by the onset detection algorithm itself.

A potential feature of the data set is that it does not represent an idealized scenario, as a significant portion of the data would be considered relatively poor quality. This, coupled with a broad range of different waveform morphologies, provides a very robust test of the performance of the algorithm. In real world applications, poor quality data and high morphological variance are common due to the inherent challenges in TCD measurement outlined previously. As such, strong performance on perfect data is not necessarily informative as to how well an algorithm will perform in the field, so testing on data that realistically spans the spectrum of possible signals is imperative to properly evaluate performance. **Figure 1** illustrates a number of examples pulled directly from the test data set, which are meant to provide a representative survey of the range of signals encountered, both in terms of inherent signal quality and morphological variation.

For reference, **Figure 1A** shows an example of a relatively good quality signal exhibiting what we would consider a normal, healthy waveform. Several common morphological landmarks



have also been labeled for reference: the diastolic valley, dicrotic notch, first peak (P1), and second peak (P2). **Figures 1B,C** illustrate two very common types of noise found in the data set. **Figure 1B** may be the result of sampling error causing a digitized appearance, and **Figure 1C** contains a significant high frequency noise component. **Figures 1D–F** represents what one might consider to be “pathological” signal morphologies. While it is not clear that these signals represent a definitive underlying hemodynamic pathology, they nevertheless diverge in important ways from what would be considered a standard, healthy waveform and serve to demonstrate the importance of robustness to morphological variation for any onset detection algorithm. **Figure 1D** shows a typical blunted waveform as described and classified in Demchuk et al. (22), in which independent P1 and P2 are not clearly identifiable. **Figure 1E** shows another not uncommonly seen morphological variant in



which the second peak (P2) is consistently higher than the first peak (P1). **Figure 1F** depicts a signal where the dicrotic notch and diastolic valley of each beat are nearly equal, resulting in two clearly distinct peaks of different amplitudes, a characteristic that may be associated with stenotic waveforms defined in Demchuk et al. (22). **Figure 1G** is meant to be illustrative of the types of extremely poor quality signals encountered in the data, and **Figure 1H** exhibits a signal with a sharp noise spike, another common feature that might “trick” an algorithm into incorrectly identifying the start of a new beat.

The example segments in **Figure 1** are meant to be illustrative of the types of noise, signal morphologies, and quality that are present in the data set that inhibit accurate onset detection. However, these examples should not be taken to be fully representative of the entire data set. In general, the signals span a range, from relatively good signals with normal morphologies to very poor signals with pathological morphologies. This is one reason why this data set offers a very robust test of the performance of the algorithm. If the same algorithm can perform well for a wide range of signal qualities and signal morphologies, then we can be more confident that it may perform well in a real world or clinical setting where differences resulting from sonographers, patients, or equipment can all contribute to increased scan to scan variance.

### 3. METHOD

The algorithm consists of six primary steps. First, the signal is pass-band filtered to remove noise. Then, a moving difference filter (MDF), defined in Equation (1), is applied to the signal. Afterwards, the MDF signal is used to help define search windows on the CBFV signal in which to look for candidate onset locations. The beat onsets are then located within these search windows. Once all candidate onsets have been identified in this manner, a beat length analysis step is performed to identify and handle possible errors. Finally, a beat alignment step is performed to fine tune the precise locations of the onsets. An overview of this general process is illustrated in **Figure 2**. Each of these steps is described in more detail in the subsections below.

#### 3.1. Band-Pass Filter

The signal is band-pass filtered using a fourth-order Butterworth filter with a passband range of 0.5–10 Hz. This filtering is done primarily to remove high frequency noise, which can complicate the matter of finding local extrema, as well as low frequency motion or breathing artifacts. We feel that this choice of passband range preserves all of the physiologically relevant features for onset detection. This filtered signal is passed along to the subsequent steps to be used for locating onsets.

#### 3.2. Moving Difference Filter

A moving difference filter (MDF) is applied to the signal in order to enhance the sharp upslope that defines the start of a typical CBFV pulse. The MDF is defined as follows:

$$z_i = y_i - y_{i-w} \quad (1)$$

where  $w$  is the length of the analyzing window and  $y$  is the filtered CBFV signal. The MDF serves a similar purpose as the slope sum function used in Zong et al. (14), but we find that the MDF performs better in terms of filtering noise and preventing false detections. The size of the analyzing window should be roughly equal to the length of the initial upslope of a typical CBFV pulse in order to maximally enhance the upslope and be less sensitive to noise artifacts. In this work, the analyzing window is taken to be 150 ms, though the performance of the algorithm is fairly robust

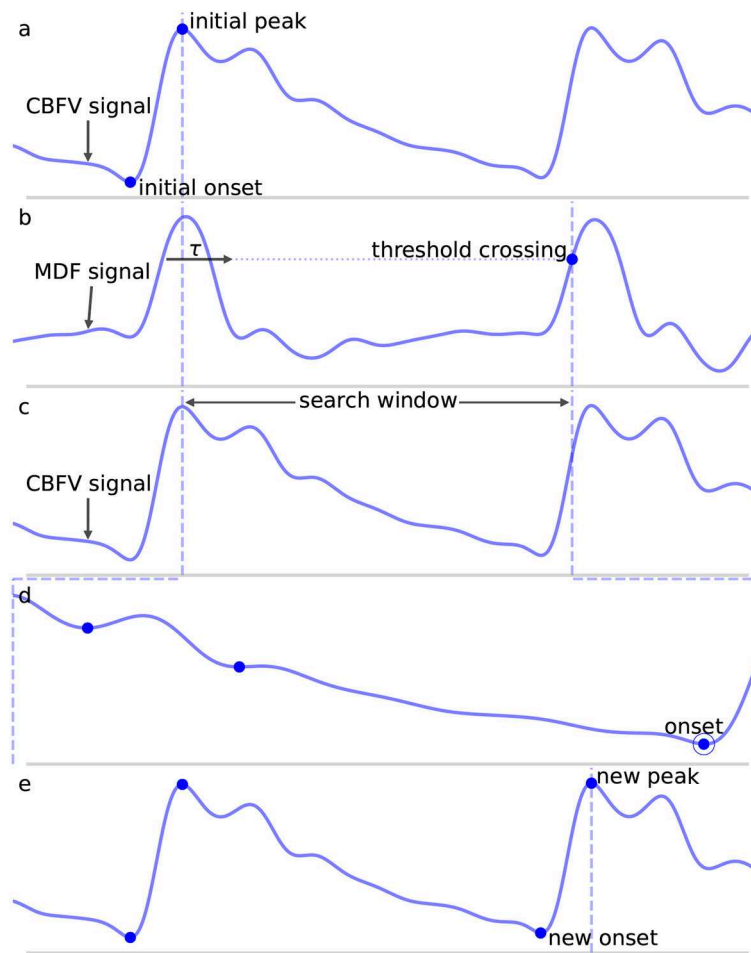
to slight changes around this value as discussed in more detail in section 4.2. An example of a CBFV waveform with a detected pulse is shown in **Figure 3a** with its corresponding MDF signal displayed in **Figure 3b**.

#### 3.3. Window Locations

Using the MDF signal, windows in which a pulse onset are likely to occur are identified. This is accomplished by adaptively thresholding on the MDF signal. The threshold is established at 60% of the average of the preceding 20 detected peaks in the MDF signal, where a peak is defined to be the local maximum immediately following the identification of a threshold crossing point. Near the beginning of a scan prior to finding 20 peaks, the threshold value is simply calculated from all the peaks that have been found so far. The justification for using a finite number of peaks (as opposed to a single peak or all peaks) is that the properties of the signal which affect the threshold value can vary on longer time scales, so thresholds should be based on peaks in the local region preceding a particular peak. However, enough peaks should be used to guard against the possibility of overweighting transient anomalies in the signal. In practice, varying this number does not have a significant impact on the algorithm's performance, as discussed in section 4.2. For the initialization step, before any peaks have been found, all peaks above 2.5 times the average of the MDF signal over the first 10 s of data are identified, and the initial threshold is set to 60% of the median value of these peaks. We find that this process generates a reasonably accurate and robust estimate of the initial threshold.

Once a threshold value has been determined, a threshold line is propagated from the previous threshold crossing point until it crosses the MDF signal, as shown in **Figure 3b** by the horizontal dotted line. The point where this line crosses the MDF signal is referred to as a threshold crossing point and is also labeled in **Figure 3b**. In order to avoid detecting multiple threshold crossing points associated with the same MDF peak, a refractory period  $\tau$  is enforced immediately following the detection of a threshold crossing point, during which no new threshold crossing points can be found. The refractory period  $\tau$  is also labeled in **Figure 3b**. The value of  $\tau$  is taken to be 200 ms, though the precise value, once again, does not have a major impact on the algorithm's performance (refer to section 4.2 for more details) so long as it is longer than the typical pulse upslope time but significantly shorter than an entire pulse, defined as the time between successive onsets.

A search window, shown in **Figure 3c** between the dashed vertical lines, is then defined as the region between the location of this threshold crossing point and the peak of the last detected pulse immediately preceding each new search window. In the case of the very first onset, for which there is no preceding peak, the search window is extended to the very beginning of the signal. The peaks of each beat should occur very close to the threshold crossing point, and are determined by finding the maximum value which occurs within the refractory interval  $\tau$ . We note here that the algorithm does not explicitly set out to accurately identify systolic peaks, and the finding of peaks is only performed in order to facilitate defining onset search windows. Thus, no claims are made regarding the accuracy of the systolic peaks.



**FIGURE 3** | Shown are the major steps in the algorithm, from an initial seed onset to detection of the next onset. The x (time) and y (CBFV/MDF amplitude) axes have been normalized to dimensionless quantities for ease of viewing. The algorithm proceeds from onset to onset in this manner until the end of the scan is reached. **(a)** Shows the initial onset in the CBFV signal, which consists of the onset point itself and its associated peak. **(b)** Displays the MDF signal associated with this same segment of data and shows how the search window is established after enforcing the refractory period and locating the threshold crossing point. **(c)** Shows the CBFV signal with the search window overlaid. **(d)** Shows a zoom-in of the search window and the location of all valleys, with the new onset labeled. Finally, **(e)** shows the original CBFV signal with the new onset identified.

### 3.4. Onset Identification

Once a search window has been identified, all valleys (identified by finding local minima) in the CBFV signal that occur within this window are located. The latest, i.e., closest to the threshold crossing point, occurring valley for which the condition:

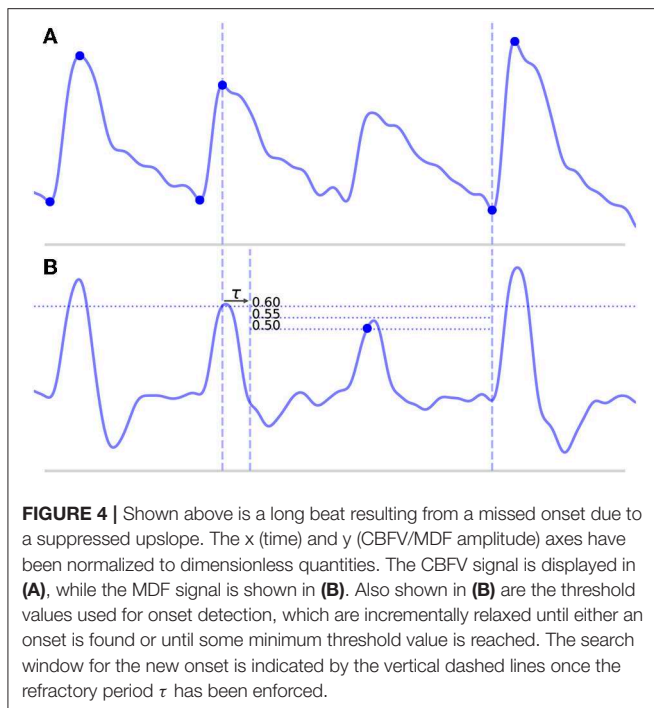
$$CBFV_{peak} - CBFV_{valley} \geq 0.75 \cdot MDF_{peak} \quad (2)$$

is satisfied is marked as the pulse onset. Here,  $CBFV_{peak}$  is the value of the peak for the CBFV pulse (the amplitude of the point labeled “new peak” in **Figure 3e**),  $CBFV_{valley}$  is the candidate onset (the points marked in **Figure 3d** with dots), and  $MDF_{peak}$  is the amplitude of the peak value of the MDF signal for this window (the value of the local maximum immediately following the threshold crossing point in **Figure 3b**). This condition (Equation 2) ensures that the algorithm avoids mistakenly identifying valleys that appear in the upslope due to noise artifacts or pathological morphologies as onsets. The constant

scale factor in front of  $MDF_{peak}$  is set to 0.75, though onset identification is fairly robust to changes in this factor (refer to section 4.2). In theory,  $CBFV_{peak} - CBFV_{valley}$  should very nearly equal  $MDF_{peak}$  since the MDF signal measures the net change in the CBFV pulse waveform over some small time window. Thus, any factor close to but slightly  $<1$  is reasonable, e.g., 0.5 – 0.95. A lower value generally provides less protection against misaligned onset detections, while too high of a value risks missing the onset altogether. The candidate onset closest to the threshold crossing point that meets the required condition is circled and labeled as “onset” in **Figure 3d**.

### 3.5. Beat Length Analysis

After the signal has been scanned through in its entirety and all initial guesses at pulse onsets have been established, the final step is to deal with outlier beats. The motivation for this is based on the fact that the initial pass tends to result in two primary kinds of mistakes.



- **Long beats** typically occur when the algorithm misses a beat, which would normally result in a false negative (FN). This most commonly occurs because of some abnormality in the upslope of a beat, either because it is not very steep and fails to cross the threshold line or because it contains some sort of noise artifact that suppresses the MDF signal. An example of such a beat is shown in **Figure 4A**, where the amplitude of a beat is suppressed such that it fails to cross the threshold line.
- **Short beats** typically occur when noise in the signal results in a sharp upslope which is erroneously identified as a new beat onset resulting in a false positive (FP). The effect of this is to divide what should be a single beat into multiple short beats.

We posit that errors of these varieties can be effectively identified by examining the distribution of beat lengths, i.e., the distance between successive onsets, resulting from the first pass of onset detection. Before proceeding however, it is worthwhile to describe a few general principles that guide a number of the choices that we make regarding the algorithm. The first principle is that we weight false negatives (failing to detect a beat onset) more severely than false positives (detecting extra beat onsets). Said another way, we believe it is better to find as many beat onsets as possible even if that means a few extra onsets are falsely identified. The motivation for this belief is influenced by the second principle, which is that the output of the algorithm should allow any downstream applications (anything using the output of the algorithm) to make the majority of the decisions about how to handle the beats once they have been segmented.

For example, in the case of short beats, these usually result from mistakenly dividing a beat into two beats, so the solution may be to just delete the onset associated with the short beat.

However, this can always be done as a post-processing step. Furthermore, such post-processing can be tailored to suit the specific needs of the downstream analysis being performed, information which is not known by the beat segmentation algorithm at runtime. Thus, as a general purpose onset detection tool, we believe it is best to be as conservative as possible, and only perform very simple post-processing steps to identify and handle the most egregious errors. This philosophy will be referenced as motivation for a number of the design decisions made in the following discussion.

### 3.5.1. Outlier Detection

In order to flag beats as outliers, we use median absolute deviation (*MAD*) due to its robustness to outliers, which was also used in Zong et al. (14). *MAD* is computed according to the following:

$$MAD = med(|X - med(X)|), \quad (3)$$

for some univariate data set  $X$  (the set of beat lengths in this case), composed of  $N$  elements  $x_1, x_2, \dots, x_N$ , where  $med$  denotes the median value. To use *MAD* as a consistent estimator of the standard deviation, a scale factor of 1.4826 is needed, which is related to the assumed normality of the distribution after excluding outliers. Thus, the estimator  $\hat{\sigma}$  for  $X$  becomes  $1.4826 \cdot MAD$ .

Finally, we define rejection criteria based on this estimator by choosing a threshold value. Setting this threshold value is a necessarily subjective choice that depends on the nature of the data and on the degree of caution deemed appropriate by the researchers. A threshold value of 3.5 has been suggested by Iglewicz and Hoaglin (23); however, a threshold value of 3.0 is also commonly suggested as a fairly conservative threshold as in Leys et al. (24). In this study, we opt to use the more conservative value of 3.5 to identify short beats and a value of 3.0 to identify long beats. We believe this to be in line with the general philosophy of performing the least amount of post-processing as possible, and to deal with only the most extreme outlier cases. The choice of an asymmetric threshold represents the goal of minimizing FN. Since long beats are generally indicative of a missed onset detection, we use a slightly less conservative threshold when identifying long beats. These choices are examined in more detail in section 4.2. Concretely, the inclusion criteria then becomes:

$$med(X) - 3.5 \cdot \hat{\sigma} < x_i < med(X) + 3.0 \cdot \hat{\sigma}. \quad (4)$$

So long as  $x_i$  satisfies (Equation 4), it is not an outlier. We can now formalize the concept of a short beat and a long beat. For a given scan for which  $N$  total beats have been detected with median length  $l_{med}$ , for a given beat with length  $l_i$  and estimator  $\hat{\sigma}$ , the beat is short if it satisfies (Equation 5) and long if it satisfies (Equation 6).

$$l_i < l_{med} - 3.5 \cdot \hat{\sigma} \quad (5)$$

$$l_i > l_{med} + 3.0 \cdot \hat{\sigma} \quad (6)$$

### 3.5.2. Long Beats

Long beats are dealt with first. The method for dealing with long beats is to apply the global onset detection algorithm to a local section of the signal, with a progressively lowered threshold line if needed. This process is detailed in the steps below. For each long beat detected:

1. The search window is defined from the beginning of the peak for the long beat until the next detected onset, as shown in **Figure 4A** and denoted by the vertical dashed lines.
2. The MDF is computed for this segment of data, as shown in **Figure 4B**.
3. A threshold line is set at 60% of the median of all the peak heights that were located in the original global MDF signal during the first pass.
4. After enforcing the refractory period, this threshold line is propagated until it either crosses the MDF signal or reaches the end of the data segment. If it crosses the MDF signal, then a search window is defined and new candidate onsets are located in exactly the same manner as described in the global onset detection algorithm.
5. If new onsets are found, these locations are saved and the algorithm proceeds to the next long beat in the scan or, if there are no more remaining long beats, the next step in the algorithm. However, if no new onsets are found, the threshold value is lowered by 0.05 and steps (3) and (4) are repeated with the lower value. The algorithm continues to lower the threshold value in this manner until either a new onset is found or some minimum threshold value is reached. A high minimum threshold value may result in more missed onsets, whereas a lower value may cause an increased number of false positives. We chose a relatively low minimum threshold value of 0.35 based on the view that false positives are preferable to false negatives (refer to section 4.2 for discussion). **Figure 4B** shows this process of new threshold lines being drawn until the missed peak is finally detected.

### 3.5.3. Short Beats

Short beats are dealt with last. This step looks at each short beat along with each of its immediate neighbors and makes a decision as to whether the short beat should be combined with either of its neighboring beats. A number of methods can be used to determine whether a merger should occur. A simple method which we find to work effectively and which is also conservative is to simply look at whether a merger of beats would produce a new beat with a length closer to the median beat length than the original neighboring beat. If so, then the short beat is examined to determine whether it looks sufficiently “different” (as defined in the following steps) from an average beat in the scan. If both conditions are met, then a merger is performed; otherwise, no merger occurs. The process for determining this is described below.

1. Four lengths are determined:  $l_{before}$ ,  $l_{short}$ ,  $l_{after}$ , and  $l_{median}$ , shown in **Figure 5**. These are defined as follows:  $l_{before}$  is the length of the beat occurring before the short beat,  $l_{short}$  is the length of the short beat,  $l_{after}$  is the length of the beat occurring after the short beat, and  $l_{median}$  is the median beat length of

all beats that have been found in the CBFV signal for a given scan. The length of a beat is defined to be the time between successive onsets.

2. The mean beat is calculated, shown in the subfigure of **Figure 5**. This is done by first finding the median beat length. All detected beats are then either truncated by removing the end or padded by repeating the last value in the beat such that all the beats have the same length, resulting in a set of length normalized vectors. The mean beat is then calculated by taking the mean of this set of vectors.
3. The algorithm checks to see if combining the short beat with the beat occurring immediately after it, resulting in a new beat length  $l_{short} + l_{after}$ , would produce a beat with a length closer to the median beat length than  $l_{after}$  alone. If it would, then one final comparison is made: the beats are combined only if the correlation distance between the short beat and the mean beat is  $>0.2$ . The correlation distance is computed by truncating the longer of the two beats such that its length is equal to that of the shorter beat, and then applying (Equation 7), where  $\mathbf{u}$  and  $\mathbf{v}$  are vectors representing each of the two beats being compared after length equalization.

$$x_{corr} = 1 - \frac{(\mathbf{u} - \bar{\mathbf{u}}) \cdot (\mathbf{v} - \bar{\mathbf{v}})}{\|(\mathbf{u} - \bar{\mathbf{u}})\| \cdot \|(\mathbf{v} - \bar{\mathbf{v}})\|} \quad (7)$$

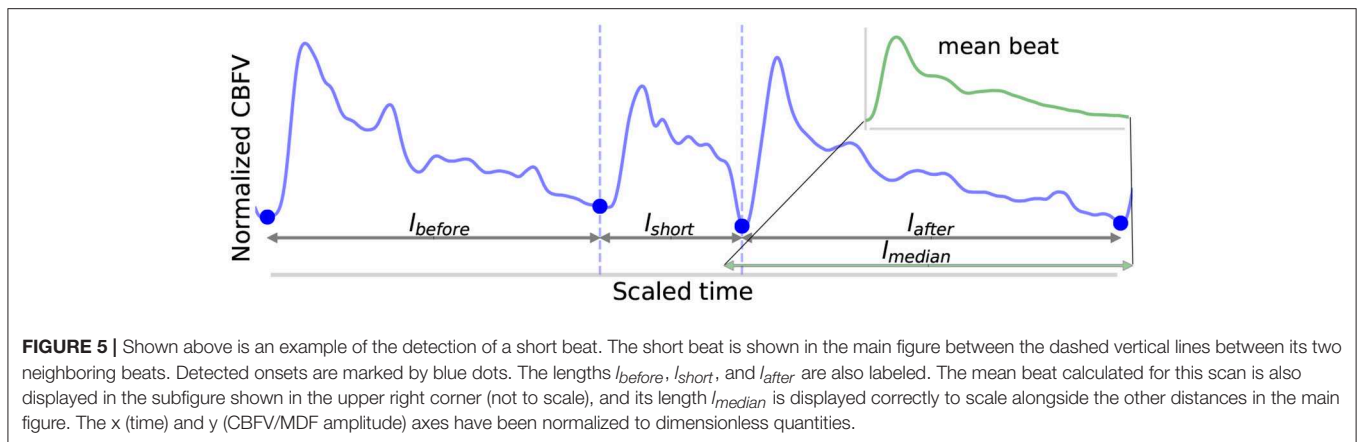
The reason for this comparison is again based on the view that false negatives are worse than false positives, and so the deletion of onsets should be done very conservatively. The motivation for enforcing this check is the idea that beats which result from noise in the signal should not look like an average beat, whereas beats that are actually just physiologically shorter will still retain the general morphological structure of a typical beat and thus should have a small correlation distance from the mean beat. If after this comparison, the beats are merged, then the algorithm moves on to the next short beat if there are any. If the beats are not merged, then the same comparison is made with the beat occurring immediately before the short beat. If no merge is performed there either, the algorithm proceeds to the next step.

4. One final step occurs in which the short beat is deleted if it is determined to be extremely different morphologically than the mean beat. In this context, “extremely” different is defined quantitatively by a correlation distance  $>0.7$ . The choice of this threshold and others is explored in more detail in section 4.2.

### 3.5.4. Iterating

It is possible for a single pass to not find all the new beats or delete all the short beats due to the fact that as beats are added/subtracted, the statistics may change very slightly. Another way to miss a beat is due to the fact that when parsing a long beat, the algorithm will exit upon finding the first new beat. If a long beat contains more than two beats, then this means a single pass will not be enough to locate all the new beats. To deal with this, the beat length analysis is performed repeatedly until no new beats are added/subtracted during a single iteration. In very rare instances (this never occurs in the data set used in this work), it





may be possible to reach an oscillating solution, so a maximum number of iterations of 10 is enforced, whereby the loop will exit regardless.

### 3.6. Beat Alignment

The final step is to fine tune the alignment of the onsets. While not strictly necessary, we find that including this step provides a significant benefit for onsets which have a large degree of misalignment relative to their annotated locations. In qualitative terms, the alignment is performed by computing a scalar value for each onset. This value represents the estimated amount by which the onset should be shifted forward or backward in time in order to better align with the actual foot position of the beat. If this estimated amount is larger than some threshold, then the shift is applied to the onset. The reason for imposing a threshold is that for large estimated shifts, we find that the shift generally moves the onset closer to its actual position. However, for small estimated shifts, corresponding to cases where the algorithm has already closely identified the actual onset, the benefit to trying to apply further realignment is questionable and in general, seems to result in slightly worse performance for very small misalignments. As a result, if the estimated amount is less than the threshold, the shift is not applied.

In order to compute the actual shift amount, the mean beat is first computed in the same way as described in section 3.5.3. The beginning portion of the beat is extracted by taking the first  $N$  samples of the mean beat, where  $N$  is equal to the number of samples in the refractory period  $\tau$ . This segment will be referred to as the mean beat upslope. The motivation for focusing on the beginning portion of the beat is the assumption that the relevant morphological feature to align on in a CBFV pulse is the initial sharp upslope. Using an extraction window equal to  $\tau$  works well as a general rule of thumb method for capturing the initial upslope along with some of the systolic peak. Next, a set of new potential onset positions are generated by taking a window of the CBFV signal and sliding it one sample at a time, starting a time  $\tau$  before the detected onset and ending a time  $\tau$  after the onset. After normalizing each vector, a dot product is then calculated between each of these new shifted onset positions and the normalized mean beat upslope. The

temporal separation between the detected onset and the shifted onset which maximizes the dot product is the estimated shift. If this shift is above 30 ms, then the shift is applied; otherwise, no action is performed. This threshold can be set by the researcher based on their individual misalignment tolerance.

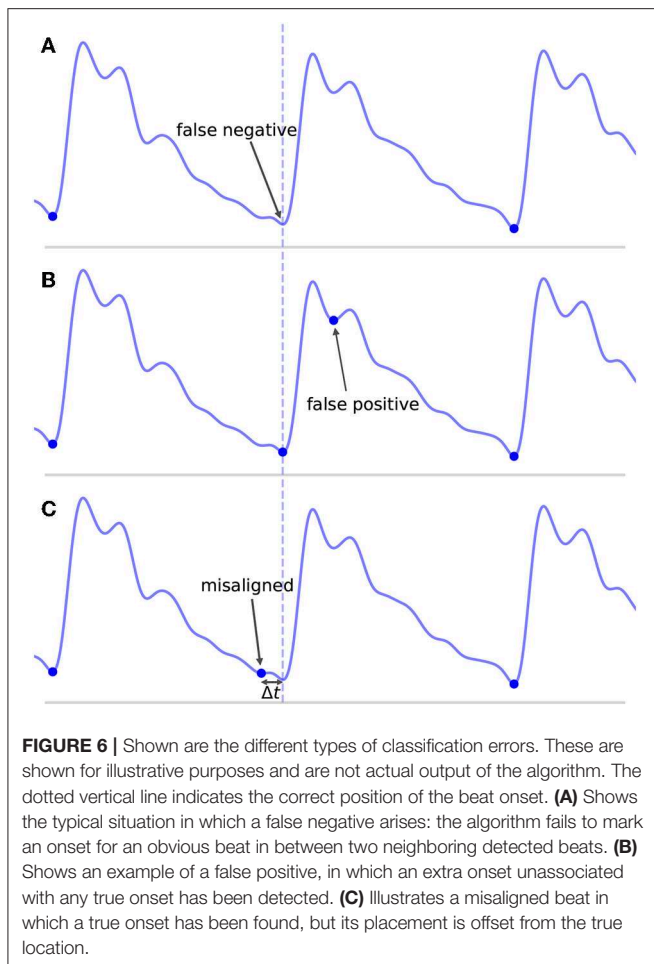
As mentioned, this step has no effect on the number of FN or FP and serves only to better align the beats in cases of large misalignments. An exploration of alternative methods for performing this alignment could be the subject of future work.

## 4. RESULTS

### 4.1. Performance Metrics

We report on a number of metrics to evaluate the performance of the algorithm. These metrics are based on the number of true positives (TP), the number of false positives, the number of false negatives, and the temporal separation between the annotated beat and the detected beat or temporal offset error ( $\Delta t$ ). For our purposes, metrics involving true negatives (TN) are not particularly meaningful since (1) almost every point is a TN and (2) the meaning of a TN becomes ambiguous in the region immediately surrounding a true onset.

Even defining what it means to be a TP, FP, and FN can prove problematic. A common convention in the literature is to define some arbitrary threshold value  $\Delta t_{max}$ . If an onset is detected such that the distance between the detected onset and the nearest true onset is less than  $\Delta t_{max}$  and there are no other closer detected onsets, then the detected onset is a TP. If no true onset exists within the threshold window or there is another detected onset closer to the true onset, then the detected onset is a FP. Similarly, for a true onset, if no detected onset exists within the threshold window, then this results in a FN. Where this can become problematic is in the case where an onset is detected in association with a true onset, but their temporal offset error  $\Delta t$  is greater than  $\Delta t_{max}$ . This situation is represented in **Figure 6C** for  $\Delta t > \Delta t_{max}$ . What is intuitively either a single mistake or potentially no mistake at all is counted as two mistakes, both a FP, because no true onset exists “near” the detected onset, and a FN, because no detected onset exists “near” the true onset. In these situations, it may be necessary to make a subjective decision



about whether to classify these types of mistakes as FP or FN. This situation is not uncommon in signals with high amounts of noise or poor signal quality or for signals whose morphology diverges significantly from what would typically constitute a normal, healthy waveform. In these cases, there may be large amounts of uncertainty when manually labeling the location of the onset, but this uncertainty is ignored when enforcing a uniform and universal classification threshold.

For these reasons, we adopt what we believe to be a more intuitive approach to error classification in the context of beat segmentation. We avoid enforcing an arbitrary threshold value and instead focus on mapping true onsets to their associated detected onsets and then quantify the temporal offset error as a separate measure. As long as a detected onset can be associated with a true onset straightforwardly in a one-to-one manner, then the pair is labeled a TP and the temporal offset error is measured. Concretely, this is done by starting with the true onsets and locating the closest detected onset, being careful to avoid double counting by ensuring that a detected onset is paired uniquely with only its closest true onset (though we note that while this situation of multiple pairings was checked for, it was never actually encountered in the results presented in this work). Any true onsets left without an associated detected onset are labeled FN. Any detected onsets left over after this

**TABLE 1** | Shown are the relevant performance metrics of the algorithm on the entire data set.

| TP    | FN | FP | TPR     | PPV     | $\overline{\Delta t}$ (ms) |
|-------|----|----|---------|---------|----------------------------|
| 92010 | 2  | 2  | 0.99998 | 0.99998 | 6.098 ± 4.749              |

process are labeled FP. A comprehensive set of unit tests were designed, and all TP pairs with large separations (>100 ms), of which there were 44 cases, as well as all FN and FP cases, were manually inspected to confirm correct behavior of the procedure.

This classification framework naturally lends itself to three types of nominal errors: FN, FP, and a new error type we will refer to as misaligned. Examples of each of these error types are shown in **Figure 6**. By classifying errors in this way, we get the following intuitive interpretations for the various types of errors: FN can be thought of as missing onsets, FP as extra onsets, and misaligned as onsets that are detected but which are offset from where they should be. The degree to which misalignment occurs is encoded by the distribution of  $\Delta t$  for all annotated-detected onset pairs.

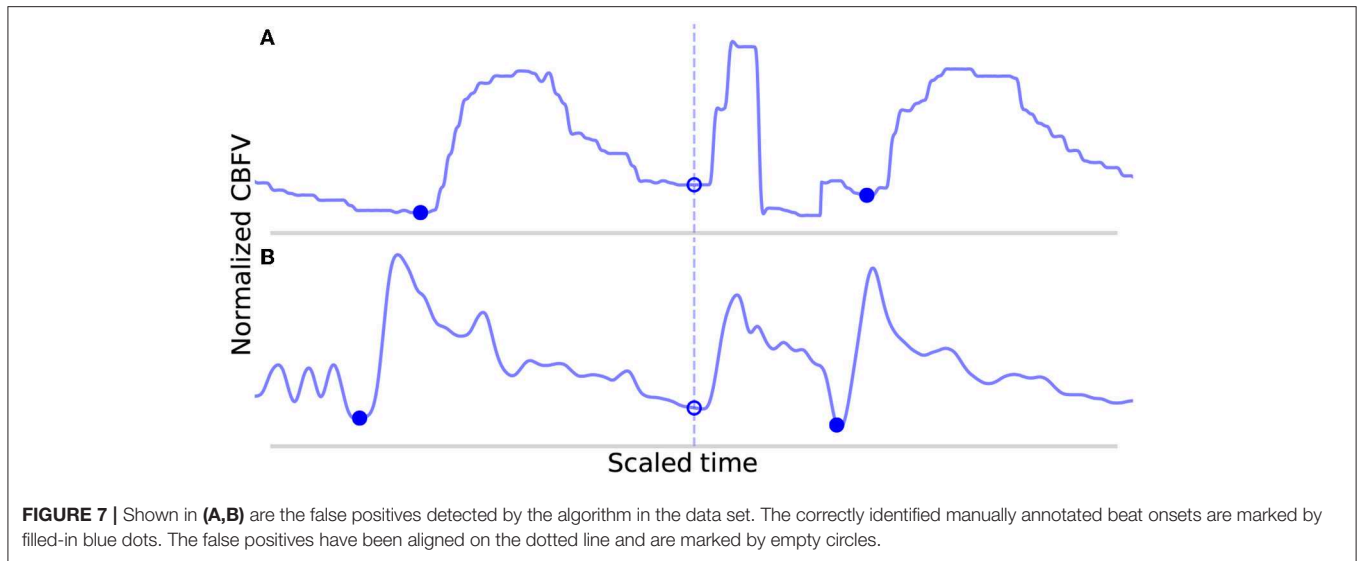
Algorithm performance is reported in **Table 1**, which shows the total number of TP, FP, and FN, as well as the true positive rate (TPR) or recall defined in Equation (8), and positive predictive value (PPV) or precision defined in Equation (9). The mean temporal offset error  $\overline{\Delta t}$  is also shown.

$$TPR = \frac{TP}{TP + FN} \quad (8)$$

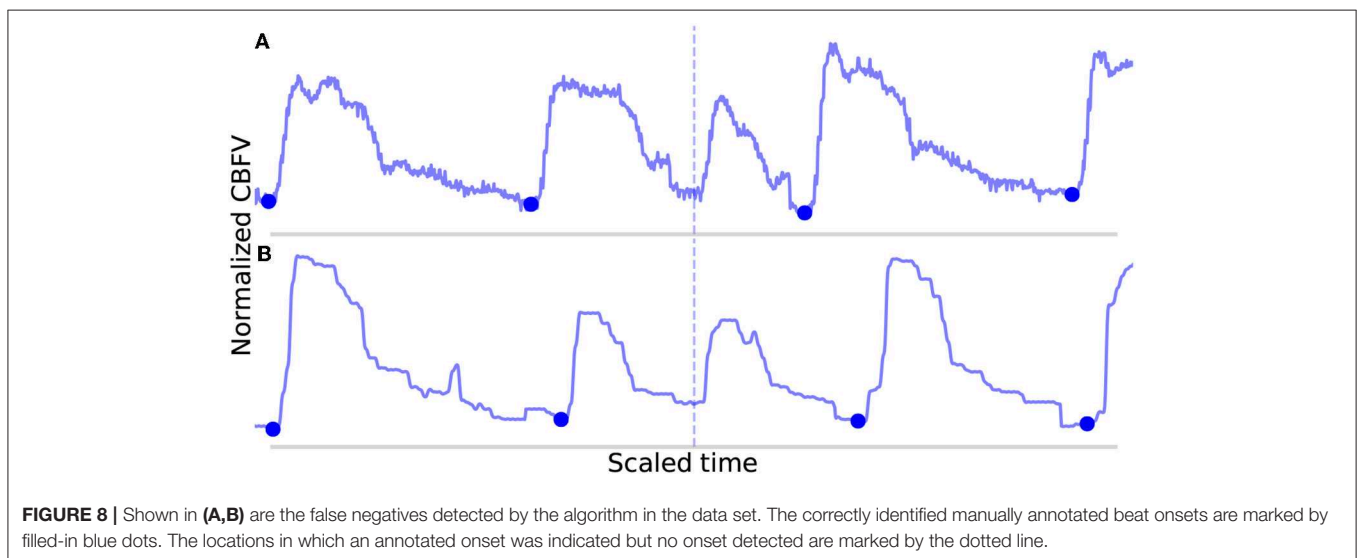
$$PPV = \frac{TP}{TP + FP} \quad (9)$$

For the data set considered in this work, the algorithm was able to detect an onset associated with nearly every annotated beat onset (2 FN) while managing to only detect two false onsets (2 FP), again given a total of 92,012 annotated onsets, for a TPR and PPV both approaching 99.998%. This represents a significant improvement over previous work using the same data set, which reported a TPR of 93.1% and a PPV of 93.3% (20). It should be noted that prior work classified TP, FP, and FN using a different convention and as a result these numbers should not be compared directly; however, the gap in performance is still sizable. Using the same 30 ms value for  $\Delta t_{max}$  as was used in Asgari et al. (20), we would obtain a TPR of 99.5% using the same error classification method as used in Asgari et al. (20).

An analysis of the failure modes of the algorithm can prove instructive into understanding the ways in which it can fail, and the combination of factors that must occur in unison to cause failures. The FP are shown in **Figure 7**. These serve to illustrate the main way in which the algorithm can falsely detect beats as both cases involve a sudden upslope in the signal likely caused by noise. Normally, these would be dealt with during the beat length analysis step, as the false detections result in short beats that would normally get merged. However, in **Figure 7A**, the FP occurs in the middle of what is an unusually



**FIGURE 7** | Shown in (A,B) are the false positives detected by the algorithm in the data set. The correctly identified manually annotated beat onsets are marked by filled-in blue dots. The false positives have been aligned on the dotted line and are marked by empty circles.



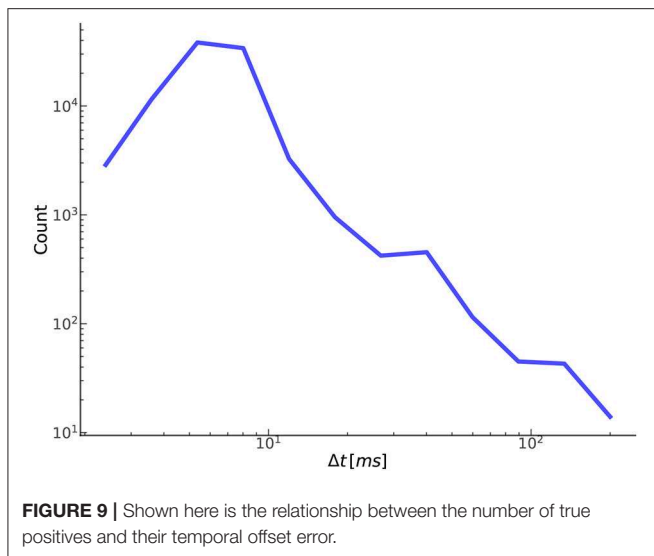
**FIGURE 8** | Shown in (A,B) are the false negatives detected by the algorithm in the data set. The correctly identified manually annotated beat onsets are marked by filled-in blue dots. The locations in which an annotated onset was indicated but no onset detected are marked by the dotted line.

long beat for this scan (the local heart rate was very low relative to the overall heart rate of this scan), and so splitting it resulted in two beats that are only slightly short and thus fail to register as short beats during outlier detection. In **Figure 7B**, the beat is identified as a short beat, but because it looks morphologically similar to a typical beat, its correlation distance from the mean beat is very low and it does not meet the criteria for merging.

The FN are shown in **Figure 8**. In **Figure 8A**, the onset is actually detected initially, but it just happens to be very short (perhaps indicative of a sudden, momentary increase in heart rate) and also just morphologically different enough from the mean beat that it registers for deletion. In **Figure 8B**, the onset is not detected during the initial pass due to the abnormally small difference between the beat's systolic peak and diastolic valley, resulting in a suppressed MDF peak. The two beats which occur here locally (the missed beat and the beat immediately preceding it) are again very short, and even when combined, fail to approach

the long beat threshold. It can be seen that the length of these two beats together is indeed very similar to the adjacent beats.

Though each of the four observed errors involved local irregularities in heart rate, the algorithm on the whole proved to be very robust to irregular heart rhythm due primarily to two key factors: the relatively short refractory period and the conservative nature of the beat length analysis procedure. The refractory period being short—which at around 200 ms, is likely shorter than any physiologically reasonable beat—means that the onset search procedure should never miss a potential onset due solely to beat length. Outlier detection with regard to beat length is inherently sensitive to heart rate irregularities; however, the conservative nature of the outlier detection appears to help guard against misidentifying excessive numbers of beats as outliers and combined with the enforced morphological criteria appears to prevent the erroneous deletion of short beats, except in the few unique cases described. Indeed, instances of irregular heart rate are both expected and commonly observed in the patient



population studied in this data set, and yet only four errors were identified in a sample of over 90,000 waveforms.

For three out of the four errors (**Figures 7B, 8**) described above, a case can be made that the true label is ambiguous based on manual inspection of the CBFV signal alone, and that perhaps one could reasonably conclude that these should not have been classified as errors. Indeed, based solely on the CBFV waveform, it is not clear why, according to the annotations, **Figure 7B** is not a true onset while **Figures 8A,B** are. However, in order to remain as objective as possible, we report only results which assume that the independently annotated labels are in fact the ground truth. Nevertheless, based on these results, it may appear that the cases in which the algorithm might fail are likely to lie at the very boundary of what would be differentiable to a human expert.

**Figure 9** shows the distribution of the temporal offset error between annotated-detected onset pairs. As expected, the vast majority of separations are  $<10$  ms with a steep dropoff thereafter. Separations on the higher side, around 50 ms or more, tend to be the result of poor quality or pathological signals, in which there is no obvious exact point of onset, but instead a region, anywhere in which an onset could reasonably be marked based on the CBFV waveform alone. Effectively, the signal quality is such that there exists a high degree of uncertainty in the start of the beat as would be judged by an expert human. **Table 2** gives the proportion of onset pairs whose temporal offset error is less than the given threshold. Already, within 10 ms, 97.8% of onsets are detected, and by 50 ms, 99.9% are detected. This table can also be interpreted as the TPR for a given  $\Delta t_{max}$  when attempting to compare results with other methods using the more conventional error classification scheme.

## 4.2. Free Parameters

A central claim in this paper is that the performance of the algorithm is not highly dependent on a large set of arbitrary and highly tuned free parameters. A case can be made that the sheer size and heterogeneity of the data set precludes the possibility of

**TABLE 2** | Shown are the proportion of annotated-detected onset pairs detected at each threshold value.

| $\Delta t_{max}$ (ms) | Proportion |
|-----------------------|------------|
| 5                     | 0.5731     |
| 10                    | 0.9778     |
| 20                    | 0.9897     |
| 30                    | 0.9951     |
| 50                    | 0.9986     |
| 100                   | 0.9995     |
| 150                   | 0.9999     |

*This can be thought of as a discretized representation of the cumulative distribution function.*

**TABLE 3** | Shown below are summary descriptions of the free parameters.

| Parameter               | Description  |
|-------------------------|--|
| Threshold normalization | Normalization factor for threshold line on MDF signal  |
| Refractory period       | Time to wait before a new onset can be detected  |
| MDF window              | Window size for computing MDF function   |
| Peaks number            | Number of peaks to include in MDF threshold line calculation                                 |
| Distance threshold 1    | Correlation distance threshold for merging beats   |
| Distance threshold 2    | Absolute correlation distance cutoff for deleting short beats                                |
| Minimum threshold norm  | Minimum normalization factor to use for threshold line when processing long beats            |
| Short outlier           | Threshold factor to use when identifying short beats   |
| Long outlier            | Threshold factor to use when identifying long beats  |
| Onset threshold         | Minimum difference between amplitude of peak and candidate onset during onset identification |

fine tuning; nevertheless, we examine here the effect of varying the free parameters on the performance of the algorithm. Though there are a sizable number of parameters, the algorithm appears to be very robust to changes in these parameters as long as they remain within reasonable ranges. These ranges are generally physiologically or mathematically motivated, though in some cases may need to be determined empirically. A summary of the free parameters along with a brief description of each parameter is presented in **Table 3**, and the recommended range for each parameter as determined in this work is presented in **Table 4**.

In order to demonstrate the algorithm's robustness, we present the magnitude of the effect of varying each parameter on the performance of the algorithm. For each parameter, a range of values are tested, and the resulting performance is recorded. Due to the low number of FN and FP relative to the total number of annotated onsets, the performance will be reported in terms of the absolute number of FN and FP, as opposed to TPR and PPV, which both remain  $\geq 0.999$  throughout the entire range of each parameter. Due to the high dimensionality of the parameter space, an exhaustive search would be difficult, so only one parameter is varied at a time, with the rest fixed to their values given in **Table 4**. We feel that this is sufficient to provide a general sense of the sensitivity of the algorithm

**TABLE 4** | Shown below are recommended ranges for each free parameter used in the algorithm along with the actual values that were used in this work.

| Parameter               | Recommended range | Value used |
|-------------------------|-------------------|------------|
| Threshold normalization | [0.5, 0.65]       | 0.6        |
| Refractory period (ms)  | [150, 300]        | 200        |
| MDF window (ms)         | [125, 200]        | 150        |
| Peaks number            | [3, 30]           | 20         |
| Distance threshold 1    | [0.05, 0.6]       | 0.2        |
| Distance threshold 2    | [0.6, 0.8]        | 0.7        |
| Minimum threshold norm  | [0.3, 0.4]        | 0.35       |
| Short outlier           | [2.5, 3.5]        | 3.5        |
| Long outlier            | [2.5, 3.5]        | 3.0        |
| Onset threshold         | [0.5, 0.9]        | 0.75       |

**TABLE 5** | Shown below are the values tested for each parameter and the resulting performance in terms of the number of FN and FP.

| Threshold norm. |          |          | Refractory period |          |          | MDF window         |          |          |
|-----------------|----------|----------|-------------------|----------|----------|--------------------|----------|----------|
| Value           | FN       | FP       | Val. (ms)         | FN       | FP       | Val. (ms)          | FN       | FP       |
| 0.5             | 1        | 11       | 100               | 2        | 9        | 110                | 3        | 8        |
| 0.55            | 2        | 4        | 150               | 2        | 8        | 130                | 3        | 3        |
| <b>0.6</b>      | <b>2</b> | <b>2</b> | <b>200</b>        | <b>2</b> | <b>2</b> | <b>150</b>         | <b>2</b> | <b>2</b> |
| 0.65            | 4        | 2        | 250               | 3        | 3        | 170                | 5        | 4        |
| 0.7             | 7        | 3        | 300               | 4        | 2        | 190                | 5        | 3        |
| Dist. thresh. 1 |          |          | Dist. thresh. 2   |          |          | Min. thresh. norm. |          |          |
| Value           | FN       | FP       | Value             | FN       | FP       | Value              | FN       | FP       |
| 0.05            | 4        | 1        | 0.5               | 3        | 2        | 0.25               | 2        | 6        |
| 0.15            | 3        | 1        | 0.6               | 2        | 2        | 0.3                | 2        | 3        |
| <b>0.2</b>      | <b>2</b> | <b>2</b> | <b>0.7</b>        | <b>2</b> | <b>2</b> | <b>0.35</b>        | <b>2</b> | <b>2</b> |
| 0.45            | 2        | 2        | 0.8               | 2        | 2        | 0.4                | 6        | 2        |
| 0.6             | 1        | 3        | 0.9               | 2        | 5        | 0.45               | 9        | 2        |
| Short outlier   |          |          | Long outlier      |          |          | Onset threshold    |          |          |
| Value           | FN       | FP       | Value             | FN       | FP       | Value              | FN       | FP       |
| 2.0             | 4        | 1        | 2.0               | 2        | 6        | 0.55               | 2        | 3        |
| 2.5             | 3        | 1        | 2.5               | 2        | 4        | 0.65               | 2        | 2        |
| 3.0             | 3        | 1        | <b>3.0</b>        | <b>2</b> | <b>2</b> | <b>0.75</b>        | <b>2</b> | <b>2</b> |
| <b>3.5</b>      | <b>2</b> | <b>2</b> | 3.5               | 4        | 2        | 0.85               | 3        | 2        |
| 4.0             | 2        | 5        | 4.0               | 5        | 2        | 0.95               | 3        | 2        |

The values that were actually used when reporting results are indicated by bold font.

to the precise value of each parameter. The results of this procedure for nine out of the ten parameters are shown in **Table 5**. The only parameter not listed is the peaks number, which showed no impact on the algorithm’s performance over the recommended range [3, 30] given the data set used in this study. It should be noted that using fewer than three peaks did negatively impact the algorithm’s performance and is therefore not recommended.

Based on these results, we conclude that the values of the free parameters primarily impact the edge cases, which by definition represent a small percentage of the total number of beats. This result was expected given the fact that many of the parameters are only involved during the beat length analysis and alignment steps, which by design were intended to catch outliers. Despite

rather large changes in some parameter values, the impact was generally negligible on overall performance. The three parameters to which the algorithm appears to be most sensitive are the threshold normalization factor, the refractory period, and the MDF window. However, the fact that a single value can perform extremely well across the wide range of subjects, waveform morphologies, and signal qualities represented in this study suggests that the parameter values established in this work may generalize well despite individual variation from scan to scan or patient to patient, without the need to perform any kind of parameter tuning.

## 5. CONCLUSION

The results presented in **Tables 1, 2** are especially promising when considering the nature of the data set. The scans were performed on patients experiencing subarachnoid hemorrhage, and the overall quality of the data generally ranges from very poor to acceptable. Numerous examples of the types of noise that may commonly be present in CBFV waveforms acquired via TCD are represented in abundance as well as a host of morphological “archetypes.” Taking into account the generally poor data quality and diverse signal morphology, we feel this data set serves as a broad collection of heterogeneous signals that may be representative of the types of data collected in real world scenarios.

Nevertheless, using our error classification framework, the algorithm presented here was able to identify all but two of the 92,012 onsets present in the data set with only two false detections. Of the correctly identified onsets, over 99.5% were detected within 30 ms of the annotated onset. These performance metrics mark a significant improvement over prior attempts at CBFV pulse onset detection and, importantly, requires no complementary signals. Additionally, the nature of the data set provides confidence that the algorithm may generalize well in a wide range of possible scenarios or deployment environments, and that it may prove useful in any signal whose pulse onsets are marked by the presence of a sharp upslope feature, as is the case in CBFV signals. More work may be warranted to show the reliability of the algorithm across all of these potential scenarios.

The algorithm’s robustness to noise and varied waveform morphologies is encouraging for its potential use in neurocritical care, where such conditions may be likely to arise during acquisition of CBFV signals. Such conditions are also likely to reduce the effectiveness of many existing onset detection algorithms, so it is important to measure the performance of any algorithm on data which exhibits these pathological qualities. It should also be noted that, while we stress the importance of the algorithm’s performance on poor quality data, the performance is expected to improve on better quality data. Indeed, none of the errors on the test set occurred in high quality signals with normal waveform morphology.

Another important consideration is the potential application of such an algorithm in real time systems. Integral to such systems is the computational complexity of the algorithm. We find that the algorithm described in this paper implemented in Python 2.7.14 is easily able to run on the entire data set in under

8 min (~5 ms per pulse) on a 2017 model Dell Precision 5520 laptop computer. Given this fact, it should be capable of running in real time on any relatively modern computer hardware with very little overhead. The only major caveat is that there are a number of steps that involve future knowledge of the signal, which would need to be modified to only use past knowledge. Given that these steps are only involved in the outlier detection steps, the impact of such changes are not likely to be significant. Quantitatively evaluating this performance impact could be the subject of future work.

## DATA AVAILABILITY STATEMENT

The datasets for this study will not be made publicly available because the academic institution does not allow public disclosure of the data.

## ETHICS STATEMENT

This study was carried out in accordance with the recommendations of the Declaration of Helsinki and the UCLA Internal Review Board. The protocol was approved by

the UCLA Internal Review Board (#10-001331). All subjects provided written and informed consent in accordance with the Declaration of Helsinki.

## AUTHOR CONTRIBUTIONS

NC, MR, MO'B, SA, FS, ST, KJ, CT, SW, and RH participated in the conception of the project, the development of various software aspects, and writing of this manuscript.

## FUNDING

This material is based upon work supported by the National Science Foundation under Grant No. 1556110. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

## ACKNOWLEDGMENTS

A draft of this manuscript has been released as a Pre-print at Canac et al. (25).

## REFERENCES

- Aaslid R, Markwalder TM, Nornes H. Noninvasive transcranial Doppler ultrasound recording of flow velocity in basal cerebral arteries. *J Neurosurg.* (1982) 57:769–74.
- Sorond FA, Hollenberg NK, Panych LP, Fisher ND. Brain blood flow and velocity. *J Ultrasound Med.* (2010) 29:1017–22. doi: 10.7863/jum.2010.29.7.1017
- Ko SB. Multimodality monitoring in the neurointensive care unit: a special perspective for patients with stroke. *J Stroke.* (2013) 15:99. doi: 10.5853/jos.2013.15.2.99
- Asgari S, Gonzalez N, Subudhi AW, Hamilton R, Vespa P, Bergsneider M, et al. Continuous detection of cerebral vasodilatation and vasoconstriction using intracranial pulse morphological template matching. *PLoS ONE.* (2012) 7:e50795. doi: 10.1371/journal.pone.0050795
- Alexandrov AV, Sloan MA, Tegeler CH, Newell DN, Lumsden A, Garami Z, et al. Practice standards for transcranial Doppler (TCD) ultrasound. Part II. Clinical indications and expected outcomes. *J Neuroimag.* (2012) 22:215–24. doi: 10.1111/j.1552-6569.2010.00523.x
- Kashif FM, Verghese GC, Novak V, Czosnyka M, Heldt T. Model-based noninvasive estimation of intracranial pressure from cerebral blood flow velocity and arterial pressure. *Sci Trans Med.* (2012) 4:129ra44. doi: 10.1126/scitranslmed.3003249
- Kim S, Scalzo F, Bergsneider M, Vespa P, Martin N, Hu X. Noninvasive intracranial pressure assessment based on a data-mining approach using a nonlinear mapping function. *IEEE Trans Biomed Eng.* (2012) 59:619–26. doi: 10.1109/TBME.2010.2093897
- Kim S, Hamilton R, Pineles S, Bergsneider M, Hu X. Noninvasive intracranial hypertension detection utilizing semisupervised learning. *IEEE Trans Biomed Eng.* (2013) 60:1126–33. doi: 10.1109/TBME.2012.2227477
- Maugans TA, Farley C, Altaye M, Leach J, Cecil KM. Pediatric sports-related concussion produces cerebral blood flow alterations. *Pediatrics.* (2012) 129:28–37. doi: 10.1542/peds.2011-2083
- Wang Y, Nelson LD, LaRoche AA, Pfaller AY, Nencka AS, Koch KM, et al. Cerebral blood flow alterations in acute sport-related concussion. *J Neurotrauma.* (2016) 33:1227–36. doi: 10.1089/neu.2015.4072
- Panerai RB. Complexity of the human cerebral circulation. *Philos Trans A Math Phys Eng Sci.* (2009) 367:1319–36. doi: 10.1098/rsta.2008.0264
- Noraky J, Verghese GC, Searls DE, Lioutas VA, Sonni S, Thomas A, et al. Noninvasive intracranial pressure determination in patients with subarachnoid hemorrhage. In: Ang B-T editor. *Intracranial Pressure and Brain Monitoring XV.* Singapore: Springer (2016). p. 65–8.
- Kazanavicius E, Gircys R, Vrubliauskas A, Lugin S. Mathematical methods for determining the foot point of the arterial pulse wave and evaluation of proposed methods. *Inform Technol Control.* (2005) 34:29–36. doi: 10.5755/j01.itc.34.1.11968
- Zong W, Heldt T, Moody G, Mark R. An open-source algorithm to detect onset of arterial blood pressure pulses. In: Murray A, editor. *Computers in Cardiology, 2003.* Thessaloniki Chalkidiki: IEEE (2003). p. 259–62. doi: 10.1007/978-3-319-22533-3-13
- Chen L, Reisner AT, Reifman J. Automated beat onset and peak detection algorithm for field-collected photoplethysmograms. In: *Engineering in Medicine and Biology Society, 2009. EMBC 2009. Annual International Conference of the IEEE.* Minneapolis, MN: IEEE (2009). p. 5689–92.
- Ferro BR, Aguilera AR, de la Vara Prieto RF. Automated detection of the onset and systolic peak in the pulse wave using Hilbert transform. *Biomed Signal Process Control.* (2015) 20:78–84. doi: 10.1016/j.bspc.2015.04.009
- Jang DG, Park S, Hahn M, Park SH. A real-time pulse peak detection algorithm for the photoplethysmogram. *Int J Electron Electr Eng.* (2014) 2:45–49. doi: 10.12720/ijeee.2.1.45-49
- Shin HS, Lee C, Lee M. Adaptive threshold method for the peak detection of photoplethysmographic waveform. *Comput Biol Med.* (2009) 39:1145–52. doi: 10.1016/j.compbiomed.2009.10.006
- Xu P, Bergsneider M, Hu X. Pulse onset detection using neighbor pulse-based signal enhancement. *Med Eng Phys.* (2009) 31:337–45. doi: 10.1016/j.medengphy.2008.06.005
- Asgari S, Arevalo NK, Hamilton R, Hanchey D, Scalzo F. Cerebral blood flow velocity pulse onset detection using adaptive thresholding. In: *2017 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI).* Orlando, FL: IEEE (2017). p. 377–80. doi: 10.1109/BHI.2017.7897284
- Asgari S, Canac N, Hamilton R, and Scalzo F. Identification of pulse onset on cerebral blood flow velocity waveforms: A comparative study. *Bio Med Res Int.* (2019). 2019:1–12. doi: 10.1155/2019/3252178

22. Demchuk AM, Christou I, Wein TH, Felberg RA, Malkoff M, Grotta JC, et al. Specific transcranial Doppler flow findings related to the presence and site of arterial occlusion. *Stroke*. (2000) 31:140–6. doi: 10.1161/01.STR.31.1.140
23. Iglewicz B, Hoaglin D. How to detect and handle outliers. In: Mykytka EF, editor. *The ASQC Basic References in Quality Control: Statistical Techniques*. Milwaukee, WI: ASQC Quality Press (1993). p. 11–13.
24. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J Exp Soc Psychol*. (2013) 49:764–6. doi: 10.1016/j.jesp.2013.03.013
25. Canac N, Ranjbaran M, O'Brien M, Asgari S, Scalzo F, Thorpe S, et al. Algorithm for Reliable Detection of Beat Onsets in Cerebral Blood Flow Velocity Signals. *bioRxiv*. (2018). p. 364356. doi: 10.1101/364356

**Conflict of Interest:** At the time that this research was conducted, NC, MR, MO'B, FS, ST, KJ, CT, SW, and RH were employees of, and either hold stock or stock options in, Neural Analytics, Inc.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Canac, Ranjbaran, O'Brien, Asgari, Scalzo, Thorpe, Jalaleddini, Thibeault, Wilk and Hamilton. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.