

Using Visual Analytics of Heart Rate Variation to Aid in Diagnostics

Stephen McIntyre, J. Mikael Eklund, and Christopher Collins

University of Ontario Institute of Technology

Abstract We present an interactive visualization tool for exploring RR interval data (the time between consecutive heart beats) to support diagnostics. An RR interval sequence diagram allows us to reduce hours of data into a general overview opposed to using short term ECG strips. A simple moving average is applied to the sequence diagram to smooth out short-term variance and highlight long-term trends. The moving average is surrounded by standard deviation bands which allow us to see the fluctuations in variance. After a brief training period using these tools coupled with RR interval and RR interval difference histograms, non-expert participants (undergraduate students) were able to differentiate between normal, atrial fibrillation, and congestive heart failure.

Keywords: HRV, Time domain, RR intervals, Information visualization, Clinical decision support system, Atrial fibrillation, Congestive heart failure

1 Introduction

Physiological signals can be complex in nature due to their tendency of being affected by multiple different aspects. Therefore, it can be harder to determine if a signal is normal than it is to determine if it is abnormal because “normal” can come in many different forms. With age and disease, patterns and predictable behaviors start to appear that allow us to diagnose specific conditions.

When working with cardiac data clinicians often refer to ECG strips, which restrict them to a small window in time that they can realistically observe. Heart Rate Variation (HRV) analysis focuses on instantaneous heart rates or RR intervals (time between consecutive heart beats). This allows exploration of longer periods of time, providing a general overview of the entire signal.

We present an interactive visualization for exploring RR interval data using multiple statistical analysis techniques to support clinical decisions. Within the domain of HRV analysis, there are many different tools [7] freely available to the public. These offer several different ways of measuring HRV either in the time or frequency domains or by non-linear methods. Our visualization expands on the current methods of analyzing RR intervals in the time domain.

2 Visualization Overview

Our visualization is comprised of two different sections, the main RR interval sequence diagram, and the RR interval histograms. Figure 1 shows the RR

sequence diagram and a seconds by hours scatter plot of each RR interval in the signal. With 24 hours of data, this sequence diagram could contain over 200,000 RR intervals causing them to be clustered together making it virtually impossible to trace between consecutive RR intervals. To address this, each RR interval point has an opacity of 0.1 allowing us to focus on the dense regions of RR intervals by reducing the distraction caused by outliers within the signals. We also provide the ability to apply a simple central moving average (mAvg) across the sequence diagram allowing us to smooth out short-term fluctuations and highlight long term-trends. The window size of the mAvg can be defined to suit the needs of the user. A larger window size will reduce the effect of short term variance highlighting the long term trend, while a smaller window size will highlight the short term variance.

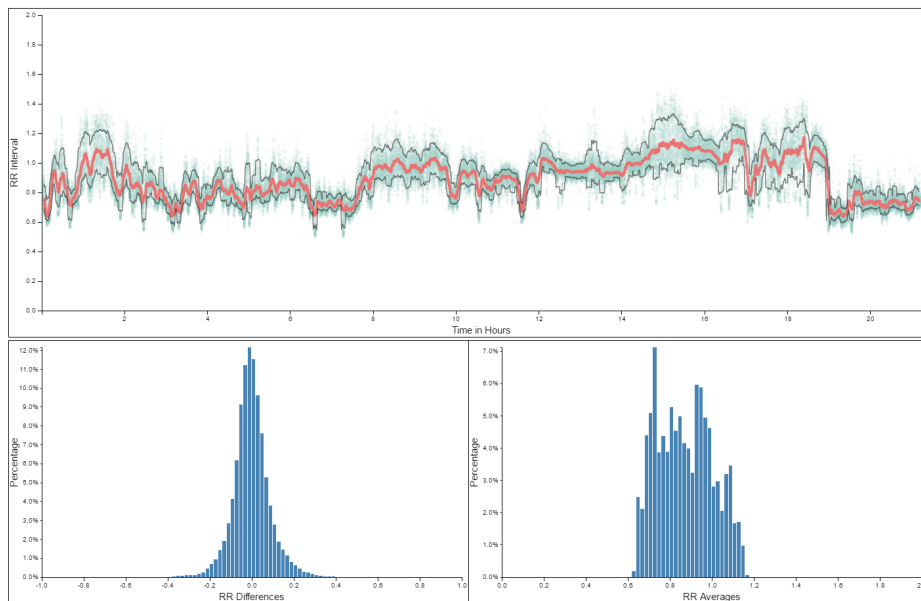


Figure 1. (Top) Main RR interval sequence diagram of a normal sinus rhythm (nsrdb 17052 [3]) with moving average in red and standard deviation bands in black. (Bottom Left) RR interval difference histogram; (Bottom Right) RR interval histogram.

The mAvg is accompanied by upper and lower standard deviation bands calculated by finding the standard deviation of the subset of values above the mean and below the mean for each average within the mAvg. The upper standard deviation is then added to the respective mean and plotted on the sequence diagram while the lower one is subtracted from the respective mean. By plotting these upper and lower deviations for each average within the mAvg, we obtain two bands that surround the mAvg. The amount of variation that exists throughout the sequence can be inferred from a quick glance at the width of the bands. A larger band represents a greater amount of variance than a thinner band.

Two histograms are included, an RR *interval histogram* (IH) and an RR *interval differences histogram* (IDH) (see Fig. 1). Significant amounts of data from a signal can be encoded into these normalized histograms, allowing easy comparison across time scales. For example a large amount of data (24 hour signal) can be compared to a small amount (4 hours) because we care more about the shape (distribution) of the histograms than we do the total values. It has been shown that when comparing or classifying signals, using both the IH and IDH combined provides more reliable qualitative information than the IH alone [2]. The IH displays the spread of the heart rates around the mean value while the IDH shows the smoothness of the rate changes. Our visualization expands on this concept by incorporating the mAvg into the creation of each histogram. During exploration, the RR sequence diagram can be brushed to select a particular period of interest within the signal. The histograms will then update to represent the selected subset of data. Depending on whether or not a mAvg has been applied to the signal, the histograms will represent either the RR average intervals or just the RR intervals themselves. Without a mAvg the IDH will show the differences between each consecutive RR interval allowing us to see the immediate short term variance. With a mAvg it will display the differences between the RR intervals and the respective mean within the mAvg, generalizing the IDH which lowers the amount short term variance can affect the final distribution.

3 Use Case Scenario

A decreased amount of short term variability has been found to be one of the indicators of congestive heart failure [6]. By looking at the RR interval differences histogram in Fig. 2 we can see that it has a thin shape. The RR interval histogram is wider because it shows us the total distribution throughout the whole signal — the long term variation. The IDH shows the difference of each consecutive RR interval which isn't affected by the overall baseline shifts throughout the signal and allows us to see the short term variance. Therefore, considering low short term variation is an indicator of CHF, the IDH may be useful in diagnosing CHF as the short term variation can be directly inferred from the chart.

By adding a moving average and standard deviation bands to the CHF signal (Fig. 3) we obtain a figure that we would expect to see from a signal with low short term variation. The short term variation causes the bands to be incredibly tight around the moving average. Considering the RR interval points are semi-transparent, these tight bands allow us to confirm that there are no clusters of RR intervals surrounding the moving average as they would increase the width of the bands. Such clusters would increase the amount of short term variance lowering the probability of CHF. It should also be noted that when adding a moving average the IDH is calculated as differences from the moving average, which is why the IDH in Fig. 3 is wider than the IDH in Fig. 2.

Therefore by using both techniques together we can infer a low short term variability from the IDH when no moving average has been applied and that no

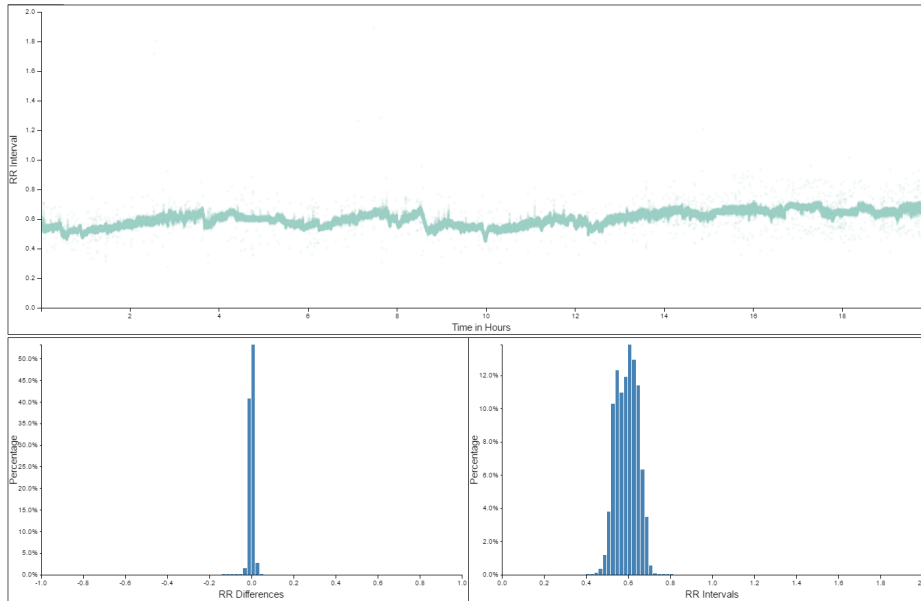


Figure 2. RR sequence diagram for congestive heart failure signal (datasets chfdb, chf05 [3,1]).

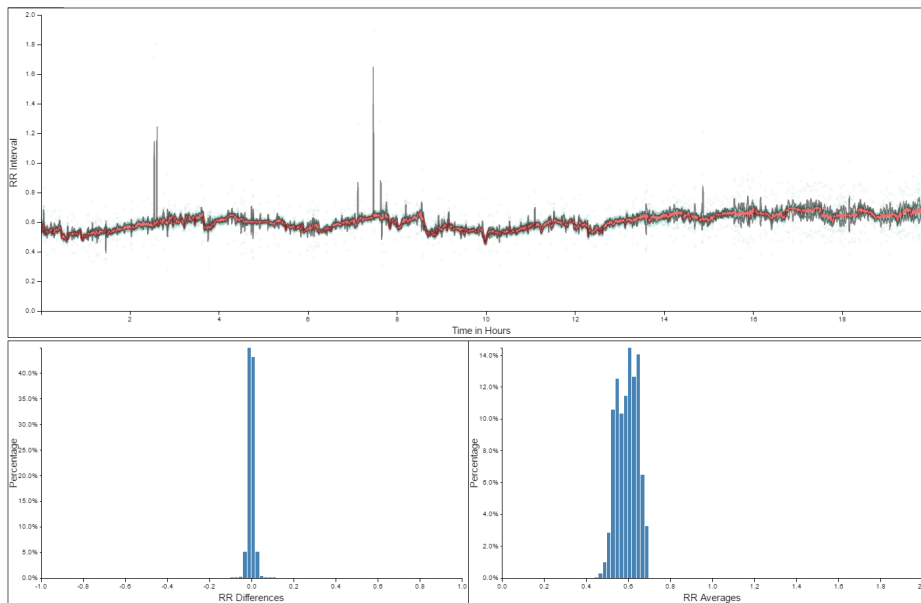


Figure 3. Moving average bands with congestive heart failure signal (datasets chfdb, chf05 [3,1]).

hidden clusters surround the moving average once the bands have been applied. Both of these are evidence that CHF could be the cause of such cardiac patterns.

Atrial fibrillation characteristically appears in the RR interval sequence as a curtain of points (essentially a large amount of short term variation). Atrial fibrillation can appear for short or long periods of time and it can be almost impossible to determine a baseline average at a glance during a period of atrial fibrillation. Without the added moving average it has been shown that atrial fibrillation produces wide triangular histogram patterns within the IDH [2]. This can be seen in Fig. 4: notice how a specific region within the signal that has been speculated to be an instance of AF has been brushed to analyze the histograms for that section.

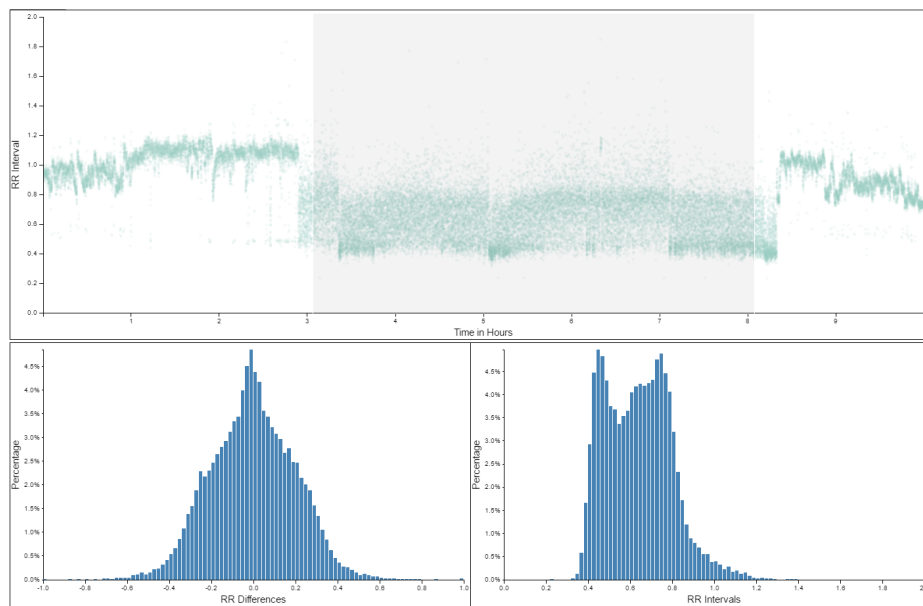


Figure 4. RR sequence diagram for atrial fibrillation signal (datasets afdb 04746 [3,5])

The moving average after being applied to the AF signal allows us to see the baseline throughout the period of AF as shown in Fig. 5. Notice how the standard deviation bands around the moving average during the period of atrial fibrillation are wide, showing that there is a high amount of short term variation. The IDH also produces a poly-modal distribution during the period of AF. Not all instances of AF have been shown to produce these poly-modal distributions but (after analyzing all the signals in the nsrdb [3]) we did not find a case where a normal sinus rhythm has produced a distribution such as this. So it may be inferred that if the IDH has a poly-modal distribution is after adding the moving average, the signal is not from the normal sinus rhythm database.

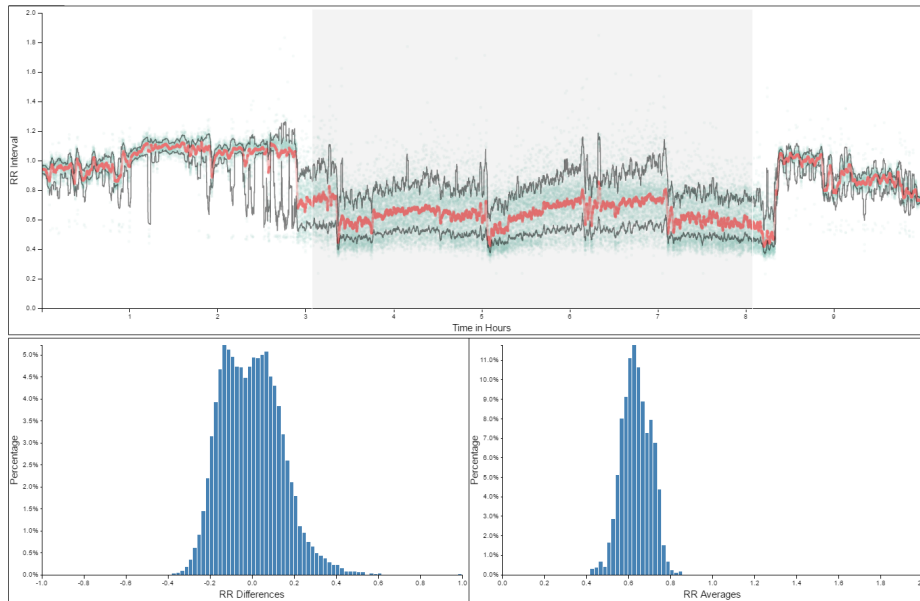


Figure 5. Moving average bands with atrial fibrillation signal (datasets afdb 04746 [3,5]) .

4 Evaluation

We performed a counterbalanced within subjects study to observe if 20 non-expert participant students would be able to accurately diagnose signals as normal or abnormal.

4.1 Data

Training and testing datasets were drawn from several data bases within the PhysioNet catalogue [3]. Each dataset was a patient ECG record lasting a minimum of 4 hours (10 hours on average). Records were collected for congestive heart failure (CHF) [1], atrial fibrillation (AF) [5], unknown cause abnormal heart rate (Abnormal) [4], and normal sinus rhythm (Normal) [3] were included. 10 testing datasets were randomly selected from each dataset, as well as 4 training datasets for each of CHF, AF, and Normal.

4.2 Task & Participants

Participants were not experts in heart rate analysis or experienced with interactive visual analytics. Participants were all undergraduate students studying in a range of programs, from health science to business and information technology.

The main factor varied in the experiment was interface style (2 levels): Basic style, and Statistical style. Participants were trained on how to use each style

and how to find characteristics of atrial fibrillation (AF), congestive heart failure (CHF), and normal sinus rhythm using that style. After hands-on training, participants were able to explore 6 training datasets provided for each style to practice making a diagnosis. During this stage participants were aware of the proper diagnosis for each signal provided and the exploration was intended for them to identify patterns within the signals they could recognize in the trial datasets.

After training they were tasked with diagnosing datasets as normal or abnormal. If they believed it to be abnormal, they also stated whether any instances of AF or CHF exist. After selecting their diagnosis they were asked to rate the confidence of their diagnosis on a scale of 1 to 7, 1 being not confident at all and 7 being completely confident. The Basic style was created to replicate the time domain analysis tools provided by a state-of-the-art HRV analysis toolkit [8]. The Basic style tasked participants with diagnosing signals while only having access to the RR interval sequence diagram and the IH. The Statistical style included everything that was available in the Basic style as well as the moving average, standard deviation bands, and the IDH.

Table 1. Confusion Matrix for Statistical and Basic Styles

Actual \ Guess	Statistical					Basic				
	Normal	Abnormal	AF	CHF	Total	Normal	Abnormal	AF	CHF	Total
Normal	86	0	7	7	100	79	4	9	8	100
Abnormal	24	20	22	34	100	22	30	20	28	100
AF	0	1	89	10	100	0	2	83	15	100
CHF	7	7	2	84	100	4	5	3	88	100
Total	117	28	120	135	400	105	41	115	139	400

4.3 Results & Discussion

Table 1 shows the diagnoses made by participants using each style. “Abnormal” signals came from the Sudden Cardiac Death Holter Database [3] which did not provide details about specific pathologies. Thus these were meant to be simply diagnosed as “Abnormal”. Table 1 shows us that participants, when using the Statistical style, were more likely to diagnosis an “Abnormal” signal as AF or CHF. The additional tools may have provided details indicative of AF or CHF causing them to diagnose as such, but more investigation is needed. While overall results were similar across techniques, during post-study interviews, participants reported that they felt more confident diagnosing a signal using the Statistical style and that they found the IDH particularly useful.

5 Future Work

One of the greatest strengths of visualizing RR interval data is that we can get an overview of an extended period. To bridge the connection between a long-term general overview and short-term ECG strips, we plan to incorporate the ECG strip into the visualization. Furthermore, discovering where to investigate the short-term data is a challenge. We plan a classifier to determine which areas of the ECG signal are of interest. These areas of interest would be highlighted to aid in the exploration of the signal, allowing faster drill down to the ECG data.

6 Conclusion

Given the success of the participants in differentiating between normal and abnormal signals after a brief training period, it can be said that patterns can be easily recognized when searching for AF or CHF. Assuming these conditions are not the only ones with easily recognizable characteristics, our visualization will be able to support clinical decisions as patterns of other conditions are discovered. A promising condition that could be investigated would be diabetic autonomic neuropathy due to the low variation of heart rate caused by this condition [2].

Acknowledgements

This research was funded by grants from NSERC.

References

1. Baim, D.S., Colucci, W.S., Monrad, E.S., Smith, H.S., Wright, R.F., Lanoue, A., Gauthier, D.F., Ransil, B.J., Grossman, W., Braunwald, E.: Survival of patients with severe congestive heart failure treated with oral milrinone. *J American College of Cardiology* 7(3), 661–670 (Mar 1986)
2. Cashman, P.: The use of RR interval and difference histograms in classifying disorders of sinus rhythm: Original articles. *Journal of Medical Engineering & Technology* 1(1), 20–28 (1977)
3. Goldberger, A.L., Amaral, L.A.N., Glass, L., Hausdorff, J.M., Ivanov, P.C., Mark, R.G., Mietus, J.E., Moody, G.B., Peng, C.K., Stanley, H.E.: PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* 101(23), e215–e220 (2000)
4. Greenwald, S.D.: Development and analysis of a ventricular fibrillation detector. Master’s thesis, Massachusetts Institute of Technology (1986)
5. Moody, G.B., Mark, R.G.: A new method for detecting atrial fibrillation using r-r intervals. *Computers in Cardiology* 10, 227–230 (1983)
6. van Ravenswaaij-Arts, C.M., Kollee, L.A., Hopman, J.C., Stoeltinga, G.B., van Geijn, H.P.: Heart rate variability. *Annals of internal medicine* 118(6), 436–447 (1993)
7. Singh, B., Bharti, N.: Software tools for heart rate variability analysis. *International Journal of Recent Scientific Research* 6(4), 3501–3506 (2015)
8. Tarvainen, M.P., Niskanen, J.P., Lipponen, J.A., Ranta-Aho, P.O., Karjalainen, P.A.: Kubios hrv–heart rate variability analysis software. *Computer Methods and Programs in Biomedicine* 113(1), 210–220 (2014)