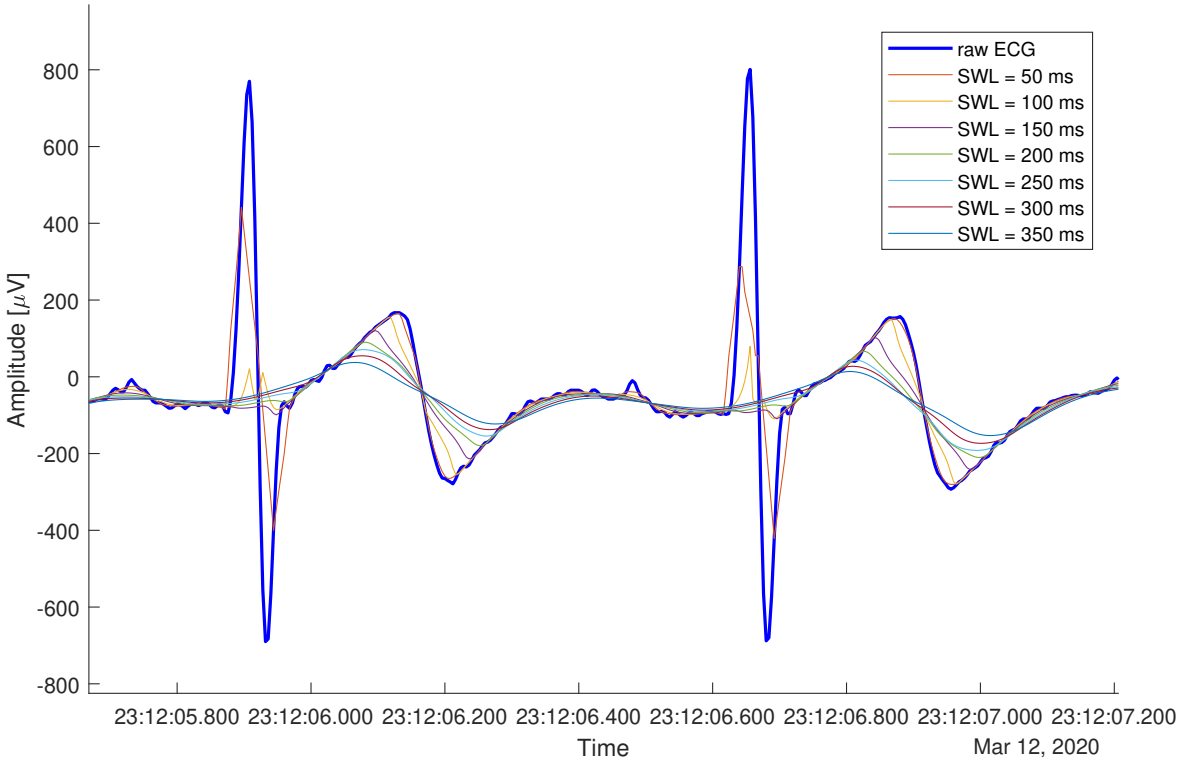
# A Appendix

## A.1 Smoothed ECG Curve

In order to preprocess the ECG signal, a smoothed ECG curve is subracted from the raw ECG signal, as described in Section A.2. The smoothed curve should be designed to closely follow the baseline and the waves of the PQRST complex, except the R wave. This was achieved using a robust locally weighted point wise linear regression over a moving time window, as described in the documentation of MATLAB[1]. A variety of smoothing window lengths (SWL) were tested to find the best length that consistently excludes the R waves while maintaining a close adherence to the rest of the curve (Supplementary Figure 1). From testing, it was determined that a SWL of 200 ms was the most suitable to meet these requirements.




**Supplementary Figure 1.** The raw ECG signal and its corresponding smoothed curves with different smoothing window lengths (SWL).

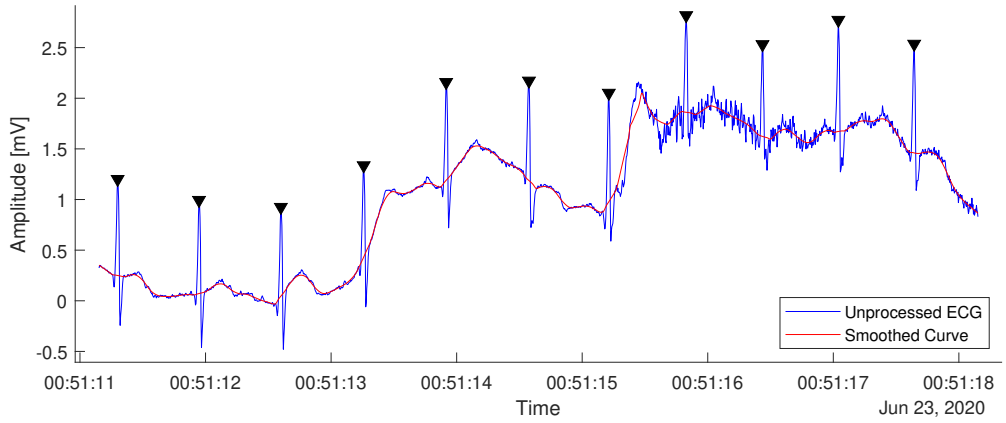
## A.2 R Wave Detection Algorithm

The proposed method to detect R waves in an ECG signal consists of three main stages: preprocessing, R wave detection, and adaptive filtering of the detected R waves.

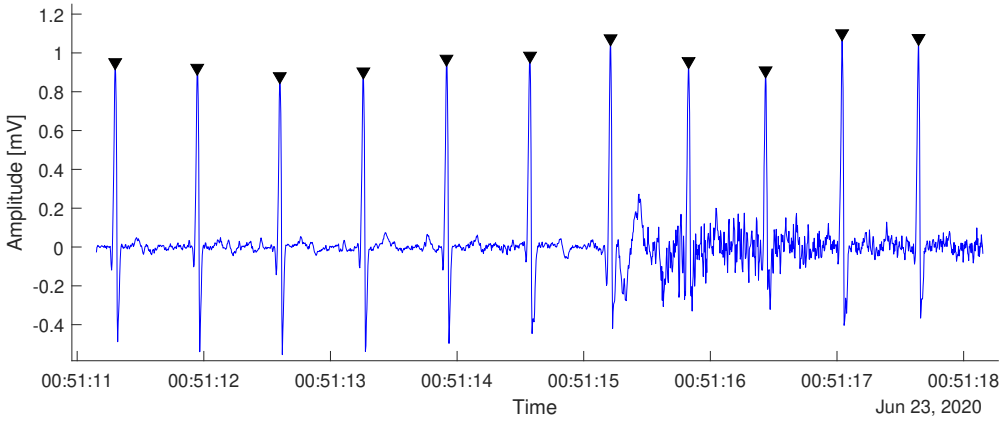
1. **Preprocessing**: In the first step, the ECG signal is preprocessed to suppress all components of the signal except for the R-waves. This is achieved by calculating a smoothed version of the signal using a robust locally weighted point wise linear regression over a moving time window with a length of 200 ms (justified in Appendix A.1)[1]. The resulting smoothed curve excludes the R-waves but follows the curves throughout the other parts of the PQRST complex. The smoothed curve is then subtracted from the original signal, resulting in a preprocessed curve in which the R-waves are emphasized while preserving their amplitudes (Supplementary Figure 2).

2. **R wave detection:** Supposed R waves in the preprocessed signal with a minimum peak height of 0.2 mV (typical R wave voltage across different leads is above 0.5 mV[2]) and a minimum peak distance of 100 ms are detected with a peak-detection function. To avoid missing any R waves that may be surrounded by other peaks, the distance between peaks is set to a implausibly small value. In cases where multiple peaks are detected within 100 ms, the most prominent peak is selected[3].

3. **Adaptive filtering of the detected R waves:** The supposed R waves are filtered by considering their widths and amplitude. The R wave height is defined as the absolute peak amplitude of the wave in Volts in the preprocessed signal. The width of the R wave is defined as the time-based width of the wave in the raw input signal at a distance of 66% of the wave height below the detected peak. Supposed R waves with a width of more than 50 ms are discarded (further explanations in Appendix A.3). From the remaining peaks, a local minimum amplitude threshold is established for each peak in a moving window of 1-minute duration. The threshold is set at 50% of the amplitude of the median of the 60 largest peaks in the window that is centered around the considered peak. Any peaks that fall below the local amplitude threshold are discarded.




**(a)** The smoothed curve (red) follows the unprocessed ECG signal (blue) but not the R waves.




**(b)** The preprocessed curve is generated by subtracting the smoothed curve from the ECG signal. This process results in suppression of changes in baseline and of the waves from the PQRST complex except the R wave.

**Supplementary Figure 2.** Illustration of the preprocessing and R wave detection procedure. The detected R waves are marked with black triangles. The amplitudes of the R waves relative to the smoothed curve are preserved.

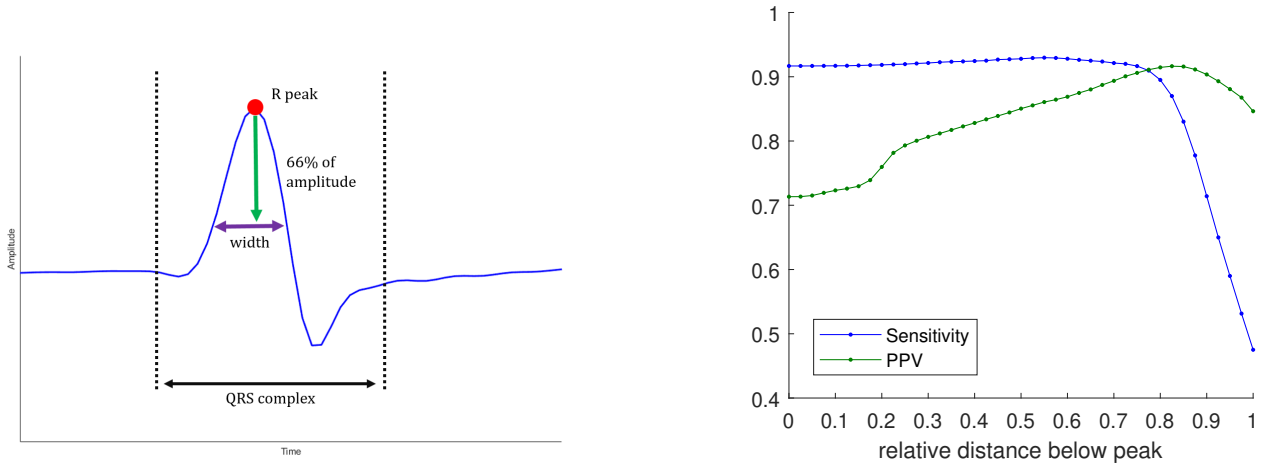
### A.3 Measurement of R Wave Width

The R wave width was measured 66% below its peak amplitude, as depicted in Supplementary Figure 3a. Measuring too close to the base may result in a lack of boundaries on either side of the curve, leading to too large or infinite R wave widths. On the other hand, measuring too close to the peak may result in a pointy artifact being incorrectly interpreted as a narrow R wave, despite its lower parts being wider than an actual R wave. In order to take into account these two effects, a threshold of 66% was chosen.

The impact of the location of the width measurement site relative to the R peak was analyzed using the MIT-BIH Noise Stress Test Database. The relative distance was varied from 0 to 1 with increments of 0.025, where 0 corresponds to the measurement of the width at the peak (always resulting in a width of 0), and 1 represents measurement of the width at the full wave amplitude below the peak. The sensitivity (Se) and the positive predictive value (PPV) were calculated. The results were

averaged across all the lead 1 data from the database (all SNR levels of records 118 and 119), for each relative distance. The results are depicted in Supplementary Figure 3b.

Normal QRS complex duration in adults is estimated to be around 80 ms, with a maximum normal duration of 100 ms[2,4]. Moreover, the R wave peak time should not exceed 45 ms in any lead[5]. Based on these characteristics, a realistic width for the R wave was determined to not exceed 50 ms at the measurement site.




**(a)** Display of a QRS complex. The width of the R wave (purple arrow) is measured at 66% of the R wave amplitude (green arrow) below the R peak (red dot).

**(b)** The relative distance below the peaks where the R wave width is measured is varied. The corresponding sensitivity and the positive predictive value (PPV) are depicted.

**Supplementary Figure 3.** Site of R wave width measurement and its effect on Sensitivity and PPV.

## A.4 Validation of R wave detection

The validity of the R wave detection algorithm was assessed using the MIT-BIH Noise Stress Test Database and only the signal from the first lead was used[6,7]. Since the proposed R wave detection algorithm (described in Section A.2) is designed to detect normal R waves, the premature ventricular contraction were excluded from the reference annotations provided with the dataset. The performance of the algorithm was evaluated based on sensitivity (Se) and positive predictive value (PPV) with respect to a tolerance of 150 ms (a clinical standard[8,9]). True positive refers to a correct detection of an R wave, false positive refers to a artefact falsely detected as an R wave and false negative refers to an R wave that was not detected.

## A.5 Validation of R wave detection

The results of the R wave detection algorithm's evaluation using the MIT-BIH Noise Stress Test Database are presented in Supplementary Supplementary Table 1. The average sensitivity and positive predictive value (PPV) for both records (118 and 119) are listed for the given signal-to-noise ratios (SNR).

**Supplementary Table 1.** Sensitivity and positive predictive value (PPV) of the proposed algorithm on different levels of signal-to-noise ratio (SNR) from the data of the MIT-BIH Noise Stress Test Database.

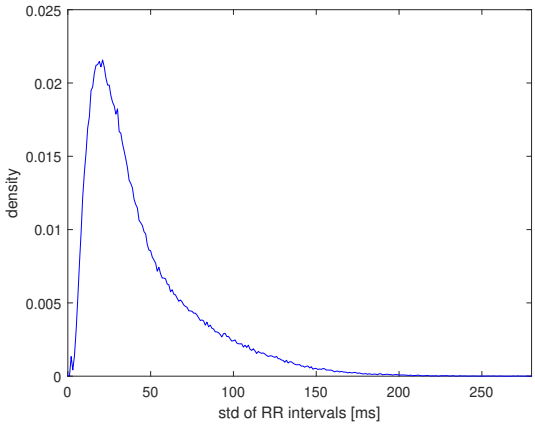
| | **SNR [dB]** | | | | | |
|---|---|---|---|---|---|---|
| | 24 | 18 | 12 | 6 | 0 | -6 |
| Sensitivity [%] | 99.61 | 99.47 | 98.53 | 93.61 | 85.47 | 73.43 |
| PPV [%] | 99.63 | 99.55 | 96.70 | 88.02 | 78.12 | 67.41 |

## A.6 Classification Thresholds and Criteria

### *A.6.1 Plausibility Criteria for RR intervals within a Segment*

The criteria to determine the plausibility of the detected RR intervals is described in Section **??**. The distribution of the standard deviation (std) of the RR intervals of the 15-second segments is displayed in Supplementary Figure 4. The analysis only considered segments where the mean and standard deviation of the RR intervals were equal between both measurement

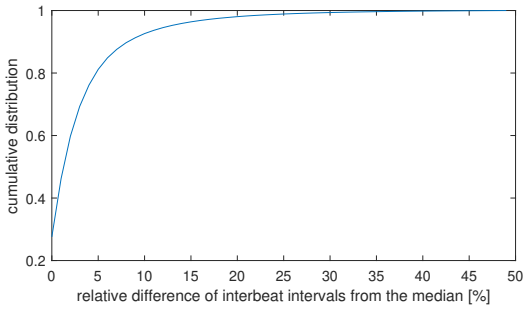
channels and therefore assumed to be correct (as outlined in Section **??**). The results indicate that the fraction of std's above 200 ms is close to zero, with the cumulative distribution between 0 and 200 ms being 0.9982 and the cumulative distribution from 200 ms to infinity being 0.0018. This highlights that it is highly unlikely for the std of the RR intervals within a 15-second segment to exceed 200 ms. Furthermore, it is physiologically impossible for the std to be Zero.




**Supplementary Figure 4.** Distribution of the standard deviations (std) of the RR intervals within the segments. The std's are binned with a bin width of 1 ms and the distribution is normalized.

#### *A.6.2 Outlier Criteria for single RR intervals*

In the proposed algorithm, the plausibility of RR intervals is assessed by comparing them to the median RR intervals within their respective 15-second segment. The cumulative distribution of the relative differences between individual RR intervals and their corresponding median is illustrated in Supplementary Figure 5. The underlying RR interval dataset comprises 4,454,000 intervals and is sourced from Dataset 1 (Section **??**). Only 15-second segments where both gel and textile electrodes measured identical RR intervals (according to the criteria outlined in Section A.7) were considered. A threshold of 33% was established for the difference between an RR interval and its corresponding median RR interval within the segment. Within the analyzed reference data, 99.2% of the RR intervals fall below this threshold.



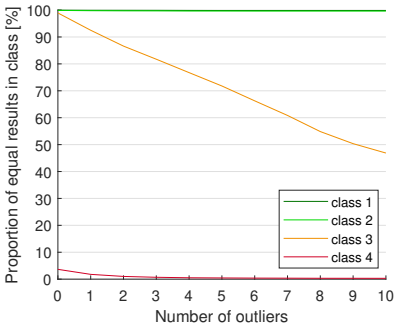

**Supplementary Figure 5.** Cumulative distribution of the relative difference of RR intervals compared to the median RR interval of the corresponding 15-second segment.

#### *A.6.3 Tolerated Number of Outliers*

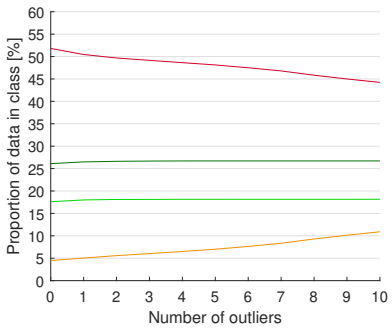
The effect of the maximum tolerated number of outliers for the plausibility assessment is depicted in Supplementary Figure 6. In the presented sensitivity (Se), specificity (TNR) and accuracy (ACC) analysis, the classification of the data is binary, with usable data (classes 1, 2, and 3) and unusable data (class 4) and was done on the reference dataset described in section **??**. True positive was defined as correctly identifying usable data (class 1, 2 and 3) and Equations 1 and 2 were used for the calculation (TP = number of true positives, FP = number of false positives, TN = number of true negatives, FN = number of false negatives). Setting the threshold at 0 results in segments being classified with high specificity but lower sensitivity, while permitting a large number of outliers results in high sensitivity but low specificity. The Accuracy peaks at 1 and 2 tolerated outliers at a value of 0.987. In the presented work, a threshold of 2 was utilized in order to prioritize the sensitivity.

$$Specifity = \frac{TN}{TN + FP} \tag{1}$$

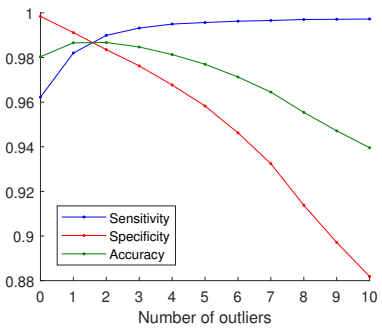
$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$




**(a)** Proportion of equal results in quality classes in relation to the threshold.




**(b)** Proportion of data in quality classes in relation to the threshold.

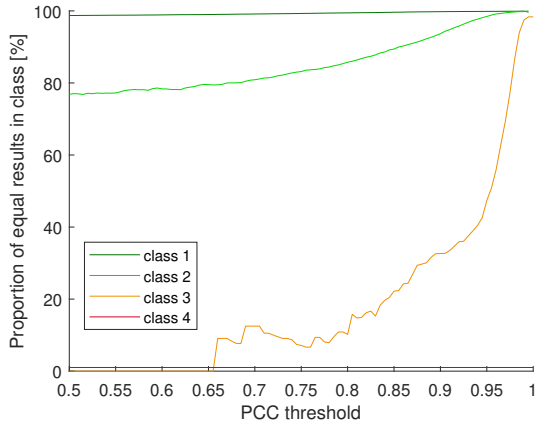



**(c)** Sensitivity, specificity and accuracy in relation to the threshold.

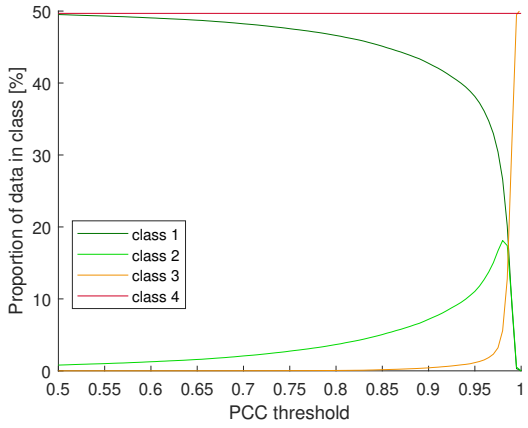
**Supplementary Figure 6.** The effect of modifying the threshold for the maximum allowable number of outliers within a segment in order to determine its plausibility.

#### *A.6.4 PCC classification threshold*

Supplementary Figure 7 shows the proportion of equal results within the quality classes and the distribution across the classes at different PCC thresholds. The analysis was done on the reference dataset decribed in section **??**. As the threshold approaches 1, the proportion of data classified as 1 and 2 diminishes. Conversely, setting the threshold too low leads to most of the data being classified as 1. The PCC threshold of 0.98 results in a well-balanced distribution of data across the quality classes.




**(a)** Proportion of equal results in quality classes in relation to PCC classification threshold.




**(b)** Proportion of data in quality classes in relation to PCC classification threshold.

**Supplementary Figure 7.** The effect of varying PCC threshold on classification.

### A.7 Criteria for equal results between two ECG channels

The maximum tolerable error in mean and standard deviation of the RR intervals in a 15 second segment is derived from the gold standard for the tolerance of QRS complex detection (150 ms[8,9]). First, it is assumed that the error of the QRS detection is normally distributed with a mean $\mu = 0$ ms. For the case of meeting the gold standard, is is assumed that 150 ms refers to the boundary of the 99% confidence interval. With Equation 3, a z-score of $z = 2.58$ and a raw score $y = 150$ ms, a standard deviation of $\sigma = 58$ ms is obtained for the error distribution. The standard error $\sigma_{\bar{x}}$ for a sample x of size n is calculated with Equation 4. With a sample size of $n = 15$ (60 bpm heart rate in 15 second window), a standard error of $\sigma_{\bar{x}} = 15$ ms is obtained. With Equation 5 and the the standard error, the 99% $ConfidenceInterval = \pm 39$ ms for the mean RR interval is obtained. For the used criteria, threshold for the standard deviation was rounded from 58 ms to 60 ms and for the mean RR interval from 39 ms to 40 ms.

$$z = \frac{y - \mu}{\sigma} \tag{3}$$

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \tag{4}$$

$$ConfidenceInterval = \mu \pm z * \sigma_{\bar{x}} \tag{5}$$

## References


1. MathWorks. *Signal Processing, Filtering and Smoothing Data (Documentation)* (MathWorks, 2022).
2. Park, M. K. Chapter 3 - electrocardiography. In Park, M. K. (ed.) *Pediatric Cardiology for Practitioners (Fifth Edition)*, 40–65, DOI: https://doi.org/10.1016/B978-0-323-04636-7.50008-3 (Mosby, Philadelphia, 2008), fifth edition edn.
3. MathWorks. *Signal Processing, Prominence (Documentation)* (MathWorks, 2022).
4. Meek, S. & Morris, F. Introduction. ii—basic terminology. *Bmj* **324**, 470–473 (2002).
5. Pérez-Riera, A. R., de Abreu, L. C., Barbosa-Barros, R., Nikus, K. C. & Baranchuk, A. R-peak time: An electrocardiographic parameter with multiple clinical applications. *Annals Noninvasive Electrocardiol.* **21**, 10–19 (2016).
6. Goldberger, A. L. *et al.* Physiobank, physiotoolkit, and physionet. *Circulation* **101**, e215–e220, DOI: 10.1161/01.CIR.101.23.e215 (2000).
7. Moody, G. B., Muldrow, W. & Mark, R. G. A noise stress test for arrhythmia detectors. *Comput. cardiology* **11**, 381–384 (1984).
8. Vollmer, M. Noise resistance of several top-scored heart beat detectors. In *2017 Computing in Cardiology (CinC)*, 1–4 (IEEE, 2017).
9. Heryan, K. *et al.* Sensitivity of qrs detection accuracy to detector temporal resolution. In *2021 Computing in Cardiology (CinC)*, vol. 48, 1–4 (IEEE, 2021).