Supplementary materials

1. Kubios automatic artifact detection algorithm

In the automatic artefact correction algorithm, artefacts are detected from dRR series, which is a time series consisting of differences between successive RR intervals. The dRR series provides a robust way to separate ectopic and misplaced beats from the normal sinus rhythm. To separate ectopic and normal beats, time varying threshold (Th) is used. To ensure adaptation to different HRV levels, the threshold is estimated from the time varying distribution of the dRR series. For each beat, quartile deviation of the 90 surrounding beats is calculated and multiplied by factor 5.2. Beats within this range cover 99.95% of all beats if RR series is normally distributed. However, RR interval series is not often normally distributed, and thus, also some of the normal beats exceed the threshold. Therefore, decision algorithm is needed to detect artefact beats. Ectopic beats form negative-positive-negative (NPN) or positive-negative-positive (PNP) patterns to the dRR series. Similarly long beats form positive-negative (PN) and short beats negative-positive (NP) patterns to the dRR series. Only dRR segments containing these patterns are classified as artefact beats. Missed or extra beats are detected by comparing current RR value with median of the surrounding 10 RR interval values (medRR). A missed beat is detected if current RR interval (RR(i)) satisfies condition RR(i) 2 − medRR(i) < 2T h (1) and an extra beat is detected if two successive RR intervals (RR(i) and RR(i+1)) satisfy condition |RR(i) + RR(i + 1) − medRR(i)| < 2T h. (2) Detected ectopic beats are corrected by replacing corrupted RR times by interpolated RR values. Similarly, too long and too short beats are corrected by interpolating new values to the RR time series. Missed beats are corrected by adding new R-wave occurrence time and extra beats are simply corrected by removing extra R-wave detection and recalculating RR interval series

2. Netstation artifact detection specification

In Netstation, the segment will be marked bad only if the number of bad channels in the segment minus the number of channels marked bad for the entire recording is greater than 30. To prevent channels that are bad for the entire recording from being incorrectly marked as bad in a segment, the Perform Inferences section runs an algorithm that detects and discards these channels from the data. It sums up the number of bad channels per segment in the remaining data and marks the segment or sample bad if it exceeds the threshold. There is an important relationship between the number of bad channels marked and the number of bad segments derived. Specifically, the more bad-channels allowed in the "Mark channel bad in recording if bad for greater than X percent" option in the Perform Inferences section, the fewer bad segments you will have. The reason is that if a channel is borderline bad, and it is not marked bad for the entire recording, then the channel will be marked bad in many segments. Consequently, in those segments, there is an additional bad channel that could trigger the maximum number-of-bad-channels criterion.

3. Strength of Bayesian Factors

Bayesian Factor (BF) 10 gives the likelihood of the data under the alternative hypothesis divided by the likelihood of the data under the null so that BF10 values greater than 1 signal more confidence in rejecting the null hypothesis and values less than 1 signal more evidence in favour of the null. The BF01 is simply 1/BF10, that is, the likelihood of the data under the null compared to the alternative. The BFs above 1 indicate correlations for which the evidence from the current study is more likely under the hypothesis that there is a relationship between those variables in the population than not. A BF greater than 3 indicates "substantial" evidence for the study hypothesis that the two variables are correlated in the population. A BF greater than 10 indicates "strong" evidence for the study hypothesis that the two variables are correlated in the population. Prior and posterior information about the correlation summarizes how our knowledge about the unknown population correlation, in which all possible values from -1 to 1 were considered equally likely (prior), has changed as a result of information gathered in our study to put more weight on positive or negative values (posterior) (Nuzzo, 2017).