

# ARTiiFACT: a tool for heart rate artifact processing and heart rate variability analysis

Tobias Kaufmann · Stefan Sütterlin · Stefan M. Schulz · Claus Vögele

Published online: 14 May 2011  
© Psychonomic Society, Inc. 2011

**Abstract** The importance of appropriate handling of artifacts in interbeat interval (IBI) data must not be underestimated. Even a single artifact may cause unreliable heart rate variability (HRV) results. Thus, a robust artifact detection algorithm and the option for manual intervention by the researcher form key components for confident HRV analysis. Here, we present ARTiiFACT, a software tool for processing electrocardiogram and IBI data. Both automated and manual artifact detection and correction are available in a graphical user interface. In addition, ARTiiFACT includes time- and frequency-based HRV analyses and descriptive statistics, thus offering the basic tools for HRV analysis. Notably, all program steps can be executed separately and allow for data export, thus offering high flexibility and interoperability with a whole range of applications.

**Keywords** Heart rate variability · ECG artifact processing · Electrocardiogram · Analysis software

In recent years, heart rate variability (HRV) has been increasingly used in medical and psychological research as a noninvasive method for the reliable estimation of vagally mediated modulation of sinus node cardiac activity. On the one hand, reduced vagal-cardiac control (i.e., low HRV) has been demonstrated to be a major risk factor for various conditions, particularly cardiovascular disorders (Kamath et

al., 1987). On the other hand, high HRV has been shown to be a valid indicator of prefrontal inhibitory capacity involved in improved executive functions, such as attentional processing or impulse control (Appelhans & Luecken, 2006; Thayer & Brosschot, 2005), and in a variety of psychopathological conditions with symptoms of impaired behavioral and emotional regulation (Thayer & Lane, 2009; Schulz, Alpers, & Hofmann, 2008). The reciprocal interconnections of the central autonomic network (Bennarroch, 1993) further highlight the clinical relevance of HRV due to a rising number of HRV biofeedback studies (for a review, see Wheat & Larkin, 2010) and techniques available to health care institutions.

There are expensive commercial HRV analysis programs such as Nevrokard® (Nevrokard, Izola/Slovenia), and powerful freely available analysis tools, such as Kubios HRV (University of Eastern Finland, Kuopio/Finland; Niskanen, Tarvainen, Ranta-aho, & Karjalainen, 2004). Nevertheless, standardization and unification of data processing have not yet resulted in the development of freely available stand-alone software comprising all necessary steps, from processing of the electrocardiogram (ECG) to IBI, detection and correction of artifacts, as well as HRV and statistical data analysis. In the present article, we describe a program that includes all of these steps. A particular difference in comparison with existing software solutions is an implemented artifact detection algorithm that is based on the intra-individual calculation of a threshold criterion and its transparent treatment.

## Artifact handling

*Artifact detection* Assessment of HRV requires the extraction of IBI, which is usually performed by extracting R-peaks from digitized ECG data. In cases in which there is

---

T. Kaufmann (✉) · S. M. Schulz  
Department of Psychology I, University of Würzburg,  
Marcusstr. 9–11,  
97070 Würzburg, Germany  
e-mail: tobias.kaufmann@uni-wuerzburg.de

S. Sütterlin · C. Vögele  
Research Unit INSIDE, University of Luxembourg,  
Campus Walferdange, Route de Diekirch - BP2,  
7220 Walferdange, Luxembourg

no raw ECG available (e.g., when recording devices such as the POLAR® RS800 [POLAR, Kempele/Finland] provide already extracted IBIs; Nunan et al., 2009), correct detection and correction of artifacts poses a particular challenge. Movement artifacts, technical failure, and poor data quality—particularly during ambulatory monitoring—increase the need for sophisticated artifact processing. Importantly, the potentially fatal impact of even minor contamination of data with artifacts (e.g., an ectopic beat resulting in a markedly shorter IBI) cannot be stressed enough (Berntson & Stowell, 1998). Despite the high importance of the particular method used for artifact detection, these procedures are rarely reported in studies assessing HRV data. The common practice of only visual inspection of extra or missing IBIs is highly questionable, in terms of reliability and practicality. Furthermore, the use of threshold values is questionable, since this gives rise to either Type I, or, more likely, Type II errors, especially when unaccounted interindividual differences in baseline heart rate (HR) are present. We argue, therefore, that the choice of artifact processing methods is often poorly justified in the literature. Importantly, such lack of consistency impairs replication and increases the risk of arbitrary decisions, particularly by inexperienced novices following these procedures.

Berntson, Quigley, Jang, and Boysen (1990) developed an artifact detection algorithm based on individual threshold criteria of artificial beats, derived from individual IBI distributions and their estimated real (not contaminated) distribution. This procedure allows for a highly automated and reliable data inspection with large data sets. Although this algorithm can still be considered as the unrivaled state-of-the-art method for IBI-based artifact detection, to our knowledge, it has not been implemented in freely available data processing software. This algorithm would be especially useful in a freely available, stand-alone software program covering the whole range of data processing from raw ECG to HRV parameters.

*Artifact correction* Several approaches for correcting artifacts in IBI series have been suggested, including linear and cubic spline interpolation, nonlinear predictive interpolation, and exclusion of ectopy-containing data segments (Lippman, Stein, & Lerman, 1994). Although promising alternatives based on artifact original IBI data alone have been reported (e.g., Clifford & Tarassenko, 2005), linear and cubic spline interpolation are still considered standard procedures for the replacement of missing or incorrect IBI, as was suggested by Malik and Camm (1995). Kubios HRV applies cubic spline interpolation to replace missing IBIs. Therefore, running comparisons against this software should rely on the same interpolation method. It appears, therefore, appropriate to include different options for artifact treatment available within one software package.

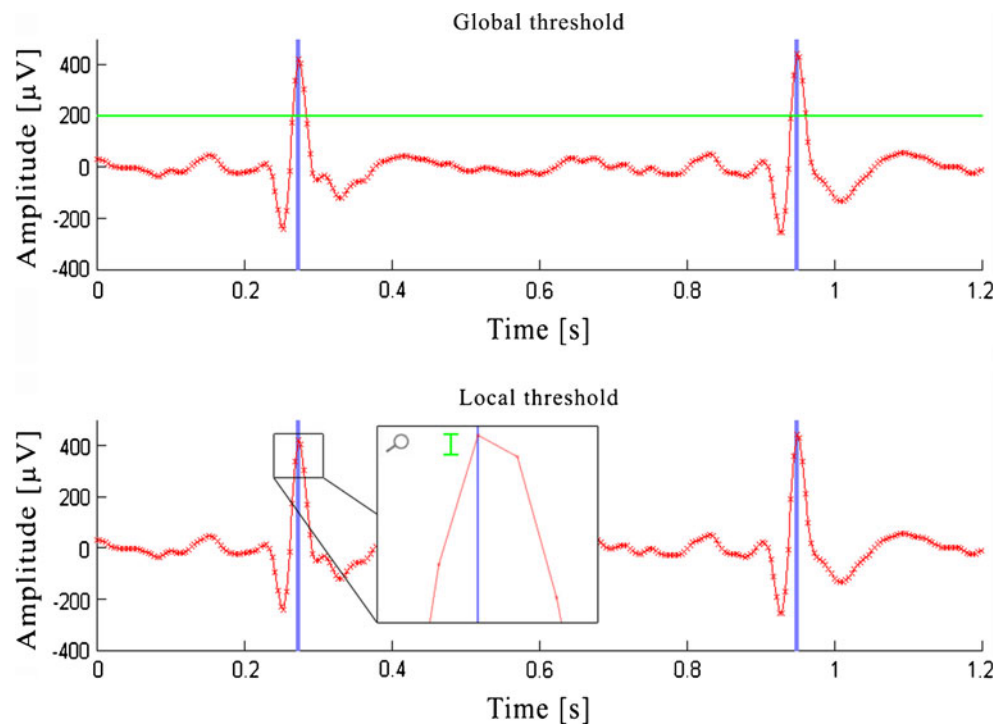
ARTiiFACT provides a powerful tool for processing of ECG and IBI data and calculation of all basic HRV parameters. In contrast with other freely available software solutions, ARTiiFACT uses independent modules allowing for various import and export opportunities and the usage of single modules only. Artifact detection is based on an automatic and intra-individually applied algorithm with excellent performance, as compared with widely used alternatives.

### Computational methods and theory

ARTiiFACT has been designed to provide researchers with a software tool covering the complete range of data processing steps, from raw ECG data to deriving HRV parameters for statistical analysis. ARTiiFACT provides (a) import options for raw ECG data as well as IBI data, (b) automated artifact detection based on distribution-related criteria determined for each individual data set, (c) interpolation of missing data (optional choice of linear or cubic spline interpolation or deletion of artifact IBI), (d) calculation of common HRV parameters in both time and frequency domains, and (e) exports for all steps of data processing via spreadsheets or text files. Notably, ARTiiFACT offers a convenient data interface to RSAtoolbox (Schulz, Ayala, Dahme, & Ritz, 2009), a freely available implementation of peak-valley analysis of RSA, including a regression-based correction of within-individual effects of breathing on estimates of vagal tone, given that the appropriate respiratory parameters (i.e., duration and volume per breath) are available (on respiratory control see e.g., Grossman & Taylor, 2007, Ritz & Dahme 2006). It is also possible to estimate breathing parameters directly from the IBI time series (e.g., O'Brien & Heneghan, 2007). This allows for partial control of respiratory effects on HRV without even recording breathing. The various I/O options of ARTiiFACT make it particularly easy to apply such corrections before submitting data to the HRV analysis module. Particular emphasis was put on intuitive handling, lean program structure, and manual intervention options. The software is MATLAB® based and is available as a stand-alone 32-bit Windows® application.

*R-peak detection* ARTiiFACT provides the possibility to low-pass or high-pass filter raw ECG data at a manually adjustable cut-off frequency. Furthermore, a window-based linear detrending method was implemented in order to purge data from long time drifts. For R-peak detection, there is a choice between global or local threshold detection criteria (see Fig. 1). Using global threshold detection, the software suggests minimum peak amplitude based on available ECG data that may be adjusted manually by the user. R-peaks exceeding this amplitude criterion are

**Fig. 1** Illustration of global and local threshold detection



detected throughout the data set. However, when drifts are present in the data, the user may switch to local threshold detection. Local thresholds are defined as the minimum voltage difference between two neighboring data points at a peak. Together, with a predefined minimum R–R distance, this allows for robust R-peak detection. If a data point exceeds the amplitude of its two surrounding points by more than the threshold criterion, and if the preceding R-peak has a greater distance than the minimum R-R distance, this value is identified as an R-peak.

*Artifact detection and treatment* The identification of spurious IBI—for example, that resulting from erroneous raw data, movement artifacts, or equipment failure—is implemented using the artifact detection algorithm proposed by Berntson et al. (1990). In the literature, manual handling of artifacts is typically based on rather subjective criteria, or, similarly, on arbitrary threshold definitions in which the artifact criteria are detected when exceeding a uniformly applied threshold (e.g., 130% above or below mean/median). Because of varying HR, both methods are not applicable across individuals and are inherently biased by existing artifacts, since they are based on data distributions, including artifacts. ARTiiFACT derives the artifact detection criterion from the distribution of IBI differences of the individual subject and applies percentile-based distribution indices, which are less sensitive to corruption by the presence of artifacts than simple threshold criteria or least-square estimates. Next, ARTiiFACT assesses the distribution parameters of an individual's IBI

dataset, removes artifacts in the first and fourth quartile, and estimates the overall (artifact-free) standard deviation on the basis of the interquartile range. On the basis of these data, the final calculation of an individual threshold criterion for beat-to-beat differences to identify artifacts is completed. Artifacts can be treated in two ways: either (a) deletion or (b) estimation, which is often done by linear or cubic spline interpolation (see Lippman et al., 1994, for a comparison of common approaches). Deleting artifacts prevents incorrect estimation of artifact IBIs but inevitably crops the data set. This reduces data reliability and may bias the data, especially when artifacts are correlated systematically with experimental conditions. In contrast, interpolation maintains both, the length and structural characteristics of the IBI series, but bears the risk of misestimating the inserted IBIs. Therefore, which of the different estimate techniques performs best may depend on the particular application.

*Time-domain methods* Commonly used time-domain parameters, such as SDNN, RMSSD, NN50, pNN50 (for a definition, see Allen, Chamber, & Tower, 2007) are calculated from the IBI time series (see Table 1).

*Frequency domain methods* Spectral frequency measures are derived using the Fast Fourier Transformation (FFT). Frequency bands are delimited in line with the Task Force's (1996) recommendations as high frequency (HF, 0.15–0.4 Hz), low frequency (LF, 0.04–0.15 Hz), and very low frequency (VLF, < .04 Hz). These frequency band definitions are default values and can be altered manually.

**Table 1** Time- and frequency-domain measures

hrvAnalysis	
Time domain	
Mean RR	Mean of interbeat intervals
Median RR	Median of interbeat intervals
SDNN	Standard deviation of interbeat intervals
RMSSD	Root mean square of successive interbeat intervals
NN50	Number of interbeat intervals that differ in more than 50 ms
pNN50	NN50 expressed as a percentage
Frequency domain	
VLF [ms <sup>2</sup> ]	Very-low-frequency component
LF [ms <sup>2</sup> ]	Low-frequency component
HF [ms <sup>2</sup> ]	High-frequency component
LF [n.u.]	Normalized units of LF
HF [n.u.]	Normalized units of HF
LF/HF	Ratio of low and high frequency

As a default, the FFT applies a Hanning window of a 256-s width, an interpolation rate of 4 Hz (spline interpolation), and an overlap of 50% to the resampled and detrended data (method of least squares). All FFT parameters can be altered manually.

*Measures of dispersion* To allow for visual evaluation of data quality and assumptions required for further data analysis (e.g., subsequent multivariate testing), standard measures of dispersion such as standard deviation, variance, and range are displayed, as well as parameters of skewness and kurtosis. The IBI distribution is shown in a histogram. Mean absolute deviation is given as a measure of dispersion less sensitive to outliers. Furthermore, the interquartile range, defined as the range between second and third quartile (Q3–Q2), and the Kolmogorov–Smirnov test for normality of the IBI distribution (alpha set to .05) are provided.

### Program description

ARTiiFACT consists of four subcomponents that are all callable from the main application window.

1. *ecgExtract*: Extraction of IBI from continuous ECG recordings.
2. *ibiArtifactProcessing*: Artifact detection and correction in IBI data.
3. *hrvAnalysis*: Calculation of HRV in time and frequency domains; see Table 1 for a list of parameters.
4. *distributionStatistics*: Tests for normality and measures of dispersion for the IBI distribution.

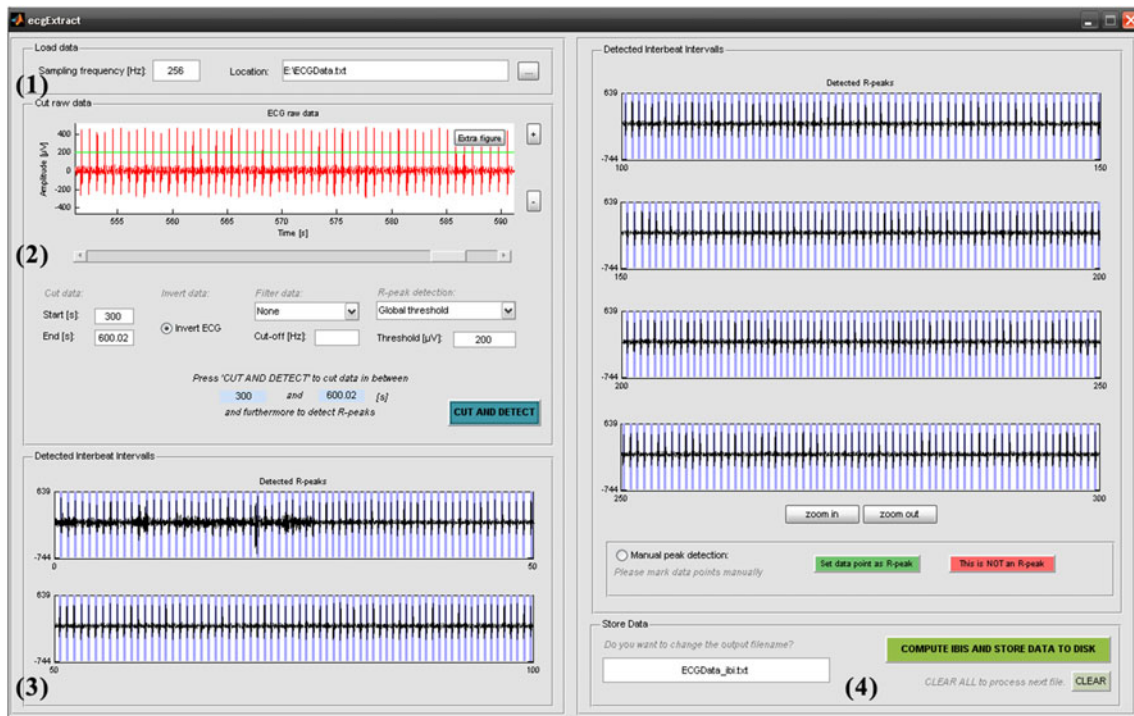
The ARTiiFACT software thus covers the full range of steps from processing raw ECG via IBI extraction to data cleaning and statistical analysis. Moreover, ARTiiFACT

enables users to exploit output from intermediate steps separately. Thus, all subcomponents can be executed independently, and their input requirements and output do not depend on or interfere with each other. For example, processing of artifacts in IBI data can be performed without any further analysis, and in this case, ARTiiFACT allows exporting artifact-corrected IBIs to a file without further processing. This program structure of independent, but seamlessly interlinked, subcomponents serves the principle of transparency and flexibility. Alternatively, users may perform a complete data analysis with ARTiiFACT using its graphical user interface (GUI). The didactic value of the structured approach to HRV analysis with this interface may establish ARTiiFACT as a reliable and transparent tool for junior researchers and in teaching contexts, as well as in medical use.

*The ecgExtract module* The subcomponent *ecgExtract* (Fig. 2) allows for import of different file formats (“\*.txt,” “\*.mat,” “\*.xls,” “\*.hdf5”) containing data of the format (channels X samples). It is possible to select the channel number (i.e., column) containing ECG data, to skip any amount of header lines, and to individually set the sampling rate. The continuous ECG data are plotted alongside a scrollable axis. Data can optionally be cut, inverted, linearly detrended, and/or filtered (high pass/low pass). A threshold for the R-peak detection can be set manually, and the appropriate detection method can be selected. R-peaks are detected and plotted for visual inspection. In case of incorrectly detected peaks, manual intervention is possible.

*The ibiArtifactProcessing module* The subprogram *ibiArtifactProcessing* requires an input text file consisting of IBI data (one column, no header), that can optionally be cut to create a suitable epoch before further processing. Detection





**Fig. 2** The graphical user interface of the subprogram ecgExtract. 1 Definition of individual sampling rate. 2 Raw ECG data plot with several options (cut, filter, invert). 3 Plot of detected R-peaks with possibility of manual intervention. 4 Data storage

of artifacts is performed as described in the section Computational Methods. It is possible to check data visually in order to detect, revise, and deselect artifacts manually. For artifact correction, there is a choice between linear or cubic spline interpolation, as well as deletion of artifact data. Interpolation replaces the detected artifacts with estimates according to the chosen interpolation method. Deletion does not insert IBI estimates, thus cropping the IBI data series. Notably, for assessing artifact position, ARTiiFACT allows exporting a file with data flags indicating the position of artifacts. All steps are displayed in the GUI (Fig. 3) and may be altered retroactively until the desired end product is achieved and saved.

*The hrvAnalysis module* HRV analysis is performed in both the time and frequency domains, which provide several highly correlated parameters indicating the extent of HRV.

*The dispersionStatistics module* Individual HRV parameters are only as accurate as the experimental conditions and accuracy of data acquisition allow. Supervision of parameters of data quality and fidelity can provide information about the quality and accuracy of the statistical outcomes and, ultimately, the validity of their interpretation. This subprogram provides descriptive measures of dispersion such as skewness, kurtosis, interquartile range, and the Kolmogorov–Smirnov test for normality, as well as quantile–quantile-plots. These data can help to assess

whether experimental conditions were appropriate and equipment was accurately used. Furthermore, it might help to identify sources of data corruption such as insufficient relaxation periods, nonstandardized stress conditions, or anxiety caused by the laboratory setting before and during data collection, which, for example, would result in a skewed distribution.

The subprograms hrvAnalysis and distributionStatistics provide the possibility to print a report sheet in the portable document format (Adobe® pdf, see Fig. 4).

### Samples of typical program runs

ARTiiFACT was tested with data sets from both ECG (recorded with Ag–AgCl electrodes according to Einthoven lead I, at 256 Hz using a g.USBamp [g.tec, Austria] amplifier) and IBI data (recorded with Polar RS800CX, Nunan et al., 2009). The following criteria were applied to these tests:

ECG raw data are sometimes contaminated with artifacts—for example, artifacts caused by movement. It appears that such artifacts are hardly distinguishable from valid R-peaks. An example (real participant data) is given in Fig. 5a, in which movement artifacts occur between seconds 8 and 10. This results in two similar peaks at  $t = 8.81$  s and  $t = 8.88$  s. In cases such as this,



**Fig. 3** The graphical user interface of the subprogram ibiArtifactProcessing. 1 Load data. 2 Raw data plot on a scrollable axis and the possibility to cut data manually. 3 Data plot with detected artifacts and

the possibility of manual intervention. 4 Data plot showing the corrected IBIs and the histogram. 5 Data storage and rsaToolbox export

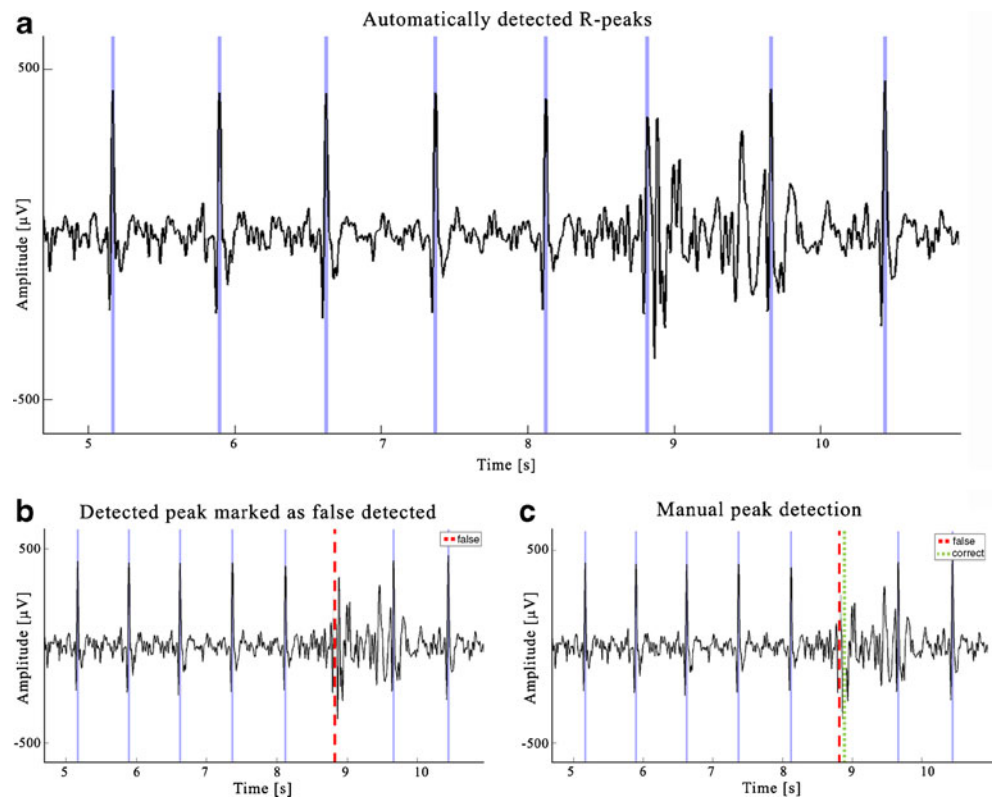
visual inspection or completely automated ECG processing are not reliable procedures to detect and distinguish

artifacts from correctly identified R-peaks. ARTiiFACT provides the possibility to deselect a detected peak



**Fig. 4** The graphical user interface of hrvAnalysis and descriptiveStatistics. 1 Load data. 2 Insert participant ID (optional). 3 Raw data plot. 4 Settings for HRV computation. 5 Time domain analysis. 6 Frequency domain analysis. 7 Store data. 8 Distribution statistics

**Fig. 5** Manual detection of R-peaks. **a** Raw data with automatically detected R-peaks. Between seconds 8 and 9, a wrong detection might have occurred, possibly due to movement artifacts in the data. ARTiiFACT gives the possibility of manual intervention, by **b** deletion of a detected R-peaks, and **c** insertion of R-peaks. However, if uncertain about the reliability of the detected R-peak, manual peak detection should not be performed, and ARTiiFACT's artifact correction should be used

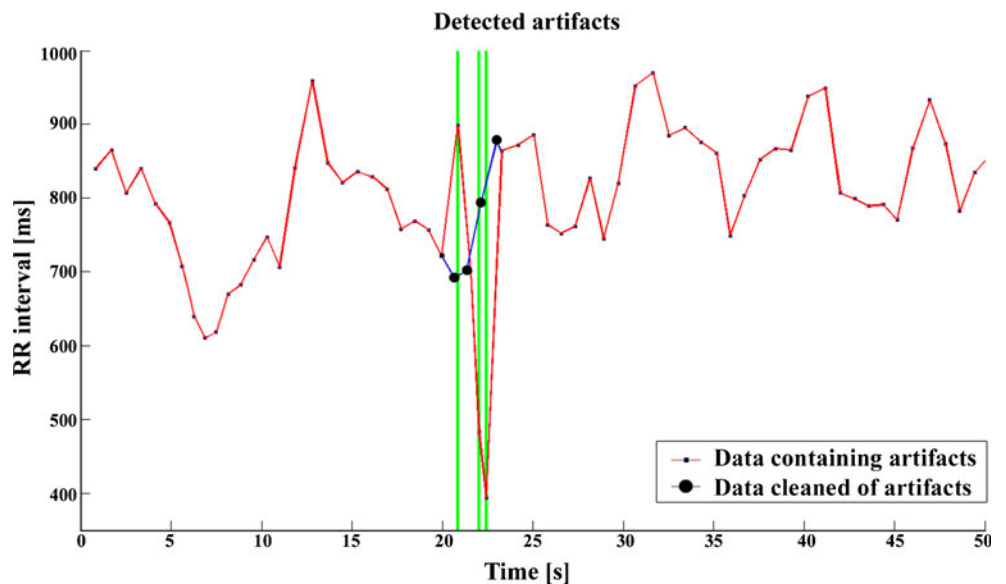


(Fig. 5b). However, if, after inspection, a peak is to be defined as true, it can be manually defined as an R-peak (Fig. 5c). Thus, the software allows fully automated artifact detection as well as complete manual control, depending on the user's preference.

Computing HRV of IBI data crucially depends on artifact-free data sets (Berntson & Stowell, 1998), since any artifact would distort variability. The algorithm recommended by Berntson et al. (1990), therefore, tries to exclude

any potential artifact with a sensitive algorithm before deriving the final criteria for identifying true artifacts. Figure 6 shows a set of IBI data containing three automatically detected artifacts. Kubios HRV applies cubic spline interpolation to replace missing IBIs. To allow for a straightforward comparison, cubic spline interpolation (blue line in Fig. 6) was used to replace artifacts when appropriate. As can be concluded from Fig. 6, the estimation fits precisely into the trend of interbeat variability.

**Fig. 6** Artifact detection and correction in IBI data. Between seconds 22 and 24, an artifact contaminates the raw data. ARTiiFACT detects all data points that are likely to be artifacts and replaces them by estimated IBIs (see text for details)



The automated artifact detection was further validated with artificial IBI data sets. These were created as follows: Starting at a mean IBI length of 833 ms, a random IBI was computed within the range of <100 ms and >5 ms in relation to the preceding IBI. This process was iterated for 100 times and resulted in a data set containing artificial IBIs. Moreover, two types of artifacts were included:

**Type A:** Two IBIs were combined to simulate a missed R-peak detection in calculating IBI data of raw ECG data. Type A IBIs are therefore about twice the length of a single typical IBI.

**Type B:** One IBI was split in half to simulate an artificial extra R-peak. Type B artifacts therefore result in two IBIs; their size can differ from equal length to major disparate lengths, which apparently is a challenge for artifact detection algorithms, since it may result in artificial intervals that are not distinguishable from valid intervals (Berntson & Stowell, 1998). For validation, different data sets were created in which the ratio of the two IBIs' sizes varied from 2.5–5% to 95–97.5%; 5–10% to 90–95%; 10–15% to 85–90%; 15–20% to 80–85%; and so on, up to 45–50% to 50–55% (the actual percentage value was chosen randomly within these constraints). This resulted in 10 different types of data sets.

Each type of data set was generated five times to avoid dependency of results on a single data set. Since each data set contained five artifacts of Type A and five artifacts of Type B, and since Type B means two artificial IBIs to detect per artifact type, 15 data points ought to have been detected in each data set.

Table 2 shows the result of this validation. In all cases, artifacts of Type A were reliably detected. For Type B artifacts, the shorter IBI was always correctly identified. As

expected, the detection of the corresponding longer IBI depended on its size. IBIs larger than 85% of the original IBI were not distinguished from valid IBIs. However, for the next 10% decrease of IBI size, detection rate increased. Only a few false alarms occurred between 75 and 90%. If the larger IBI was smaller than 75% of the actual complete IBI, all IBIs were reliably detected as artifacts.

To validate the quality of this artifact detection algorithm, the same data sets were analyzed with the Kubios HRV software, a widely used tool that is often used as HRV analysis software (see, e.g., Culbertson et al., 2010; Sütterlin, Herbert, Schmitt, Kübler, & Vögele, 2011) and offers artifact correction at five levels of sensitivity ranging from *very low* to *very strong*. Table 3 summarizes the corresponding results for the three criteria: low, medium, and strong. Both low and medium correction criteria resulted in good detection of Type A artifacts, but only poor detection of Type B artifacts. Only the shorter part of Type B artifacts was reliably detected. Unfortunately, the undetected longer part still contaminated the remaining IBI series. Although the criterion strong resulted in better detection of both parts of Type B artifacts, it did not reach ARTiiFACT's detection accuracy (see Tables 2 and 3, number of missed detections). Moreover, with the criteria set to strong, the number of false detections increased dramatically throughout the whole test data set (44 false detections in total).

ARTiiFACT, by contrast, produced a considerably lower number of false detections (12 false detections in total). Their occurrence also was restricted to test data sets where the size of larger artificial IBIs was rather large (i.e., between 85 and 90%). In sum, ARTiiFACT showed a better artifact detection rate and fewer false detections.

**Table 2** Validation of the automated artifact detection

Artifact Detection in ARTiiFACT						
Length of short IBI (% of actual IBI) for Type B artifacts	Number of Type A artifacts	Number of Type B artifacts	Number of false detections	Number of missed detections	$d'$	$\beta$
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	3	25	2,89	65,96
15–20%	25/25	31/50	2	19	3,32	90,69
20–25%	25/25	40/50	7	10	3,45	21,34
25–30%	25/25	50/50	0	0	NA	1
30–35%	25/25	50/50	0	0	NA	1
35–40%	25/25	50/50	0	0	NA	1
40–45%	25/25	50/50	0	0	NA	1

Different data sets were created, including artifacts with the whole range of possible artifact length. All data sets were validated according to the detection of two types of artifacts, either Type A or Type B artifacts. Both false alarms and true missings were validated.  $d'$  = sensitivity,  $\beta$  = response bias



**Table 3** Validation of artifact detection performance in Kubios HRV for comparison

Artifact Detection in Kubios HRV						
Length of short IBI (% of actual IBI) for Type B artifacts	Number of Type A artifacts	Number of Type B artifacts	Number of false detections	Number of missed detections	$d'$	$\beta$
<b>LOW:</b>						
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	0	25	NA	NA
15–20%	25/25	25/50	0	25	NA	NA
20–25%	25/25	25/50	0	25	NA	NA
25–30%	25/25	25/50	0	25	NA	NA
30–35%	25/25	24/50	0	26	NA	NA
35–40%	25/25	26/50	0	24	NA	NA
40–45%	25/25	31/50	0	19		
<b>MEDIUM:</b>						
2.5–5%	25/25	25/50	0	25	NA	NA
5–10%	25/25	25/50	0	25	NA	NA
10–15%	25/25	25/50	0	25	NA	NA
15–20%	25/25	25/50	0	25	NA	NA
20–25%	25/25	25/50	0	25	NA	NA
25–30%	25/25	26/50	0	24	NA	NA
30–35%	24/25	33/50	0	17	NA	NA
35–40%	25/25	47/50	0	3	NA	NA
40–45%	25/25	50/50	0	0	NA	1
<b>STRONG:</b>						
2.5–5%	25/25	25/50	6	25	2,67	35,06
5–10%	25/25	25/50	2	25	3,02	95,74
10–15%	25/25	25/50	7	25	2,57	27,28
15–20%	25/25	27/50	4	23	2,90	50,34
20–25%	25/25	34/50	7	16	3,69	160,12
25–30%	25/25	47/50	1	3	4,09	7,34
30–35%	25/25	50/50	4	0	NA	0
35–40%	25/25	50/50	10	0	NA	0
40–45%	25/25	50/50	2	0	NA	0

Kubios HRV provides artifact detection at different criteria levels ranging from very low to very strong. Results presented for the three criteria: low, medium and strong.  $d'$  = sensitivity,  $\beta$  = response bias

For further quality assessment of the artifact detection algorithm, we computed signal detection theory estimators for sensitivity ( $d'$ ) and detection bias ( $\beta$ ) (Stanislaw & Todorov, 1999). Notably, ARTiiFACT reached the goal of  $\beta = 1$  for those data sets, with the larger IBI being smaller than 75% of the actual complete IBI (see Table 2). Kubios HRV, on the other hand, reached  $\beta = 1$  only for those data sets with larger IBI smaller than 60% of the actual IBI, and only for the detection criterion medium. For the detection criteria low and strong, it failed to reach the target  $\beta$  value (see Table 3).

Therefore, we conclude that ARTiiFACT provides sensitive, well-balanced, and reliable artifact detection in IBI data, thus optimizing the accuracy of subsequent

analyses, such as computing HRV parameters (Berntson & Stowell, 1998).

### Hardware and software specifications

ARTiiFACT was developed using MATLAB® 2009b and was compiled with the MATLAB® compiler 4.13. The necessary MATLAB® component runtime 7.13 (MCR) was packaged along with the application. ARTiiFACT should work on all 32-bit Windows operating systems (XP, Vista and 7 were tested). The minimal desktop resolution is 1280 × 768 pixels.

## Mode of availability of program

ARTiiFACT is freely available upon request (email tobias.kaufmann@uni-wuerzburg.de). Additionally, the requesting user is provided with a tutorial and user manual. Registered users may opt in for notification of free updates. Current planning includes options for batch processing, extending HRV analysis by autoregressive models and nonlinear analysis, and a full module for treatment of respiration-related issues.

## Summary

In the present article, we present a software tool providing the user with an efficient artifact detection algorithm, including the possibility of manual revision, artifact removal, and computation of HRV, as well as descriptive statistics on the distribution of data. Although a broad variety of settings and functions are available, ARTiiFACT provides a graphical user interface that makes it applicable for both research and teaching purposes. Its modular structure and compatibility allows integration with other software tools, replacing one or more of ARTiiFACT's subcomponents to maximize benefits by combining advantages of various software solutions.

**Author Note** The work presented in this paper was funded by a grant from the Alfred-Krupp-Science Foundation to C. V., grant number DM8-MEDFAK 01.

## References

- Allen, B., Chambers, A. S., & Towers, D. N. (2007). The many metrics of cardiac chronotropy: A pragmatic primer and a brief comparison of metrics. *Biological Psychology*, *74*, 243–262.
- Appelhans, B. M., & Luecken, L. J. (2006). Heart rate variability as an index of regulated emotional responding. *Review of General Psychology*, *10*, 229–240.
- Bennarroch, E. E. (1993). The central autonomic network: Functional organization, dysfunction, and perspective. *Mayo Clinic Proceedings*, *68*, 988–1001.
- Berntson, G. G., Quigley, K. S., Jang, J. F., & Boysen, S. T. (1990). An approach to artifact identification: Application to heart period data. *Psychophysiology*, *27*, 586–598.
- Berntson, G. G., & Stowell, J. R. (1998). ECG artifacts and heart period variability: Don't miss a beat! *Psychophysiology*, *35*, 125–132.
- Clifford, G. D., & Tarassenko, L. (2005). Quantifying errors in spectral estimates of HRV due to beat replacement and resampling. *IEEE Transactions on Biomedical Engineering*, *52*, 630–638.
- Culbertson, C., Nicolas, S., Zaharovits, I., London, E. D., La Garza, I., Brody, A. L., et al. (2010). Methamphetamine craving induced in an online virtual reality environment. *Pharmacology, Biochemistry and Behavior*, *96*, 454–460.
- Grossman, P., & Taylor, E. W. (2007). Toward understanding respiratory sinus arrhythmia: Relations to cardiac vagal tone, evolution and biobehavioral functions. *Biological Psychology*, *74*, 263–285.
- Kamath, M. V., Ghista, D. N., Fallen, E. L., Fitchett, D., Miller, D., & McKelvie, R. (1987). Heart rate variability power spectrogram as a potential noninvasive signature of cardiac regulatory system response, mechanisms, and disorders. *Heart and Vessels*, *3*, 33–41.
- Lippman, N., Stein, K. M., & Lerman, B. B. (1994). Comparison of methods for removal of ectopy in measurement of heart rate variability. *The American Journal of Physiology*, *267*, H411–H418.
- Malik, M., & Camm, A. J. (1995). *Heart rate variability*. New York, NY: Futura.
- Niskanen, J. P., Tarvainen, M. P., Ranta-aho, P. O., & Karjalainen, P. A. (2004). Software for advanced HRV analysis. *Computer Methods and Programs in Biomedicine*, *76*, 73–81.
- Nunan, D., Donovan, G., Jakovljevic, D. G., Hodges, L., Sandercock, G. R. H., & Brodie, D. A. (2009). Validity and reliability of short-term heart-rate variability from the Polar S810. *Medicine and Science in Sports and Exercise*, *41*, 243–250.
- O'Brien, C., & Heneghan, C. (2007). A comparison of algorithms for estimation of a respiratory signal from the surface electrocardiogram. *Computers in Biology and Medicine*, *37*, 305–314.
- Ritz, T., & Dahme, B. (2006). Implementation and interpretation of respiratory sinus arrhythmia measures in psychosomatic medicine: Practice against better evidence? *Psychosomatic Medicine*, *68*, 617–627.
- Schulz, S. M., Alpers, G. W., & Hofmann, S. G. (2008). Negative self-focused cognitions mediate the effect of trait social anxiety on state anxiety. *Behavior Research and Therapy*, *46*, 438–449.
- Schulz, S. M., Ayala, E., Dahme, B., & Ritz, T. (2009). A MATLAB toolbox for correcting within-individual effects of respiration rate and tidal volume on respiratory sinus arrhythmia during variable breathing. *Behavior Research Methods*, *41*, 1121–1126.
- Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, & Computers*, *31*, 137–149.
- Sütterlin, S., Herbert, C., Schmitt, M., Kübler, A., & Vögele, C. (2011). Frames, decisions, and cardiac-autonomic control. *Social Neuroscience*, *6*, 169–177.
- Task Force of the European Society of Cardiology and The North American Society of Pacing and Electrophysiology. (1996). Heart rate variability—Standards of measurement, physiological interpretation, and clinical use. *European Heart Journal*, *17*, 354–381.
- Thayer, J. F., & Brosschot, J. F. (2005). Psychosomatics and psychopathology: Looking up and down from the brain. *Psychoneuroendocrinology*, *30*, 1050–1058.
- Thayer, J. F., & Lane, R. D. (2009). Claude Bernard and the heart-brain connection: Further elaboration of a model of neurovisceral integration. *Neuroscience and Biobehavioral Reviews*, *31*, 81–88.
- Wheat, A. L., & Larkin, K. T. (2010). Biofeedback of heart rate variability and related physiology: A critical review. *Applied Psychophysiology and Biofeedback*, *35*, 229–242.