©Journal of Sports Science and Medicine (2022) **21,** 260-266
http://www.jssm.org DOI: https://doi.org/10.52082/jssm.2022.260

Research article

RMSSD Is More Sensitive to Artifacts Than Frequency-Domain Parameters: Implication in Athletes' Monitoring

Nicolas Bourdillon [1,2]✉, Sasan Yazdani [2], Jean-Marc Vesin [3], Laurent Schmitt [4] and Grégoire P. Millet [1]
[1] Institute of Sport Sciences, University of Lausanne, 1015 Lausanne, Switzerland; [2] be.care SA, Renens, Switzerland; [3] ASPG, Applied Signal Processing Group, EPFL, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; [4] National Centre of Nordic-Ski, Research and Performance, Prémanon, France

Abstract
Easy-to-use and accurate heart rate variability (HRV) assessments are essential in athletes' follow-up, but artifacts may lead to erroneous analysis. Artifact detection and correction are the purpose of extensive literature and implemented in dedicated analysis programs. However, the effects of number and/or magnitude of artifacts on various time- or frequency-domain parameters remain unclear. The purpose of this study was to assess the effects of artifacts on HRV parameters. Root mean square of the successive differences (RMSSD), standard deviation of the normal to normal inter beat intervals (SDNN), power in the low- (LF) and high-frequency band (HF) were computed from two 4-min RR recordings in 178 participants in both supine and standing positions, respectively. RRs were modified by (1) randomly adding or subtracting 10, 30, 50 or 100 ms to the successive RRs; (2) a single artifact was manually inserted; (3) artifacts were automatically corrected from signal naturally containing artifacts. Finally, RR recordings were analyzed before and after automatic detection-correction of artifacts. Modifying each RR by 10, 30, 50 and 100 ms randomly did not significantly change HRV parameters (range -6%, +6%, supine). In contrast, by adding a single artifact, RMSSD increased by 413% and 269%, SDNN by 54% and 47% in supine and standing positions, respectively. LF and HF changed only between -3% and +8% (supine and standing) in the artifact condition. When more than 0.9% of the signal contained artifacts, RMSSD was significantly biased, whilst when more than 1.4% of the signal contained artifacts LF and HF were significantly biased. RMSSD and SDNN were more sensitive to a single artifact than LF and HF. This indicates that, when using RMSSD only, a single artifact may induce erroneous interpretation of HRV. Therefore, we recommend using both time- and frequency-domain parameters to minimize the errors in the diagnoses of health status or fatigue in athletes.

Key words: Artifact, frequency-domain, heart rate variability, noise, time-domain.

Introduction

Heart rate variability (HRV) estimates non-invasively the autonomic nervous system modulations. In athletes it has become an essential tool for monitoring fatigue, overreaching or overtraining (Plews et al., 2012; Schmitt et al., 2013, 2015). In several professional and amateur sports, from cycling to football weekly and daily follow-up are used to make critical decisions regarding training plans and competitions. It is a point of debate whether athletes' monitoring can be limited to RMSSD (Schmitt et al., 2015b) analysis or shall be performed using both the time and frequency-domain analyses (Schmitt et al., 2021).

Today, with the fast expansion of new technologies, hundreds of smartphone applications propose to monitor HRV. However, outside of the controlled conditions of a laboratory it is very difficult if not impossible to standardize recording conditions and ensure that data is of good quality (i.e., limited number of artifacts and physiological steady state) and therefore that interpretation can be made reliably.

Repeating measures on a regular basis gives insight on the true representation of individuals physiological state (Le Meur et al., 2013). Athletes repeat measures very often (Schmitt et al., 2018), whilst between one and three times per week has been proposed for long-term follow-up (Plews et al., 2012). However, HRV recordings may be noisy (Saboul and Hautier, 2019) and the reproducibility of the time- and frequency-domain parameters is not very good (interclass correlation coefficient, ICC: 0.79 and 0.57 for RMSSD and SDNN, respectively, and 0.86 and 0.47 for LF and HF, respectively) as reproducibility below 0.60 is questionable (Pitzalis et al., 1996). At the same time, the coefficient of variation in the frequency-domain remains in the 20-50% range (Sandercock et al., 2004).

It is well accepted that artifacts contribute to substantial alteration of HRV parameters and that this bias exceeds typical effect sizes seen in studies. The effects of a single artifact on a RR series are dramatic. A single artifact can increase the estimate of HF variability by almost 3 natural log units. This is a large bias relative to typical experimental effect sizes often in the range of 0.5–1.0 Ln (Berntson and Stowell, 1998). However, few studies have compared the four most commonly used HRV parameters (RMSSD, SDNN, LF and HF) with regards to their sensitivity to artifacts (Rincon et al., 2018).

There is no universal method for editing ectopic beats (Salo et al., 2001). The amount and type of RRs editing have remarkably different effects on the various HRV parameters. Editing ectopic beats corrects their effects and improves HRV stability over time (Tarkiainen et al., 2007). To date, manual assessment of recordings and artifact removal remains the gold standard, but is error prone, time consuming, and dependent on skill level and experience of the assessor (Berntson et al., 1990). For example, it is nearly impossible to perform human assessment in a football team where 35 recordings are made simultaneously in the morning and the immediately following training for each player is adapted accordingly. Smartphone applications use automated correction

Received: 01 October 2021 / Accepted: 27 April 2022 / Published (online): 1 June 2022

methods but even if they show high performances, they necessarily either leave unresolved artifacts (no correction) or modify some normal RR intervals (over-correction) (Stapelberg et al., 2018), both of which alter HRV interpretation.

Therefore, the present research questions are (1) which of the common HRV parameters are the most sensitive to artifacts; (2) how many unresolved artifacts would alter HRV interpretation.

Methods

Assessment of artifact influence on HRV parameters was performed in two steps:

1. *RR editing test*, focus was put on a set of 51 recordings (dataset 1) that naturally contained no artifact. These recording were artificially edited as detailed below. The aim of this step was to compare the effects of multiple little RR intervals modifications versus one single artifact on HRV.
2. *Automatic artifact correction*, an automatic detection-correction of artifacts was applied on a set of 178 recordings (dataset 2) that naturally contained artifacts. The aim of this step was to determine how many artifacts in a recording would alter HRV.

Dataset 1 and 2 follow the exact same recording protocol: 11-min RR recordings were collected in supine and standing positions using a chest belt (TP5, Cardiosport, Waterlooville, UK) connected via Bluetooth to the participant's smartphone. Out of the 11-min recordings (6 min supine followed by 5 min standing), the last four min in each position were processed for HRV analysis (Bourdillon et al., 2017). The RR recordings were stored in their raw format before any automated correction.

In commercial smartphone application, 4-minute windows seem unusually long but in the context of this scientific publication we focused on 4-min windows in accordance with the general recommendations of HRV analysis (Task Force, 1996; Bourdillon et al., 2017). In the present case the last 4 minutes in each position (supine and standing) was analyzed. The effects of a single artifact on HRV parameters reported in the present article on 4-minute windows would be amplified on 1-min or 30-s window duration as commonly found in many smartphone applications.

In the two datasets, the HRV parameters extracted were: the root mean square of the successive differences (RMSSD), the standard deviation of RRs (SDNN), the spectral power in the low-frequency (LF, 0.04 - 0.15 Hz) and high-frequency bands (HF, 0.15 - 0.40 Hz) in ms^2, the total power (Tot = LF + HF). The spectral power was estimated using the averaged periodogram on the resampled RR intervals (4 Hz) using a window length of 250 data points and an overlap of 50%. All computations were performed separately for the supine and standing positions using MATLAB® (MathWorks, Natick, MA, USA).

Dataset 1 was made of 51 recordings; they belonged to 51 professional and amateur athletes between 27 and 61 years old, BMI between 18 and 26 kg/m^2. All recordings were artifact free but were edited to introduce various modifications in the RR time series and therefore influence the HRV parameters.

Dataset 2 was made of 178 recordings, they belonged to 178 professional and amateur athletes between 18 and 60 years old, BMI between 18 and 25 kg/m^2. This dataset naturally contained artifacts, those artifacts were corrected and HRV parameters computed before and after correction.

RR editing test (dataset 1)

Figure 1A shows a typical 4-min RR recording in the supine position that visibly contains no artifact. This data was edited by (1) randomly adding or subtracting 10, 30, 50 or 100 ms to the successive RRs; (2) manually adding one typical artifact. These modifications are denoted 'rand10', 'rand30' 'rand50', rand100, and 'artifact' thereafter, respectively.

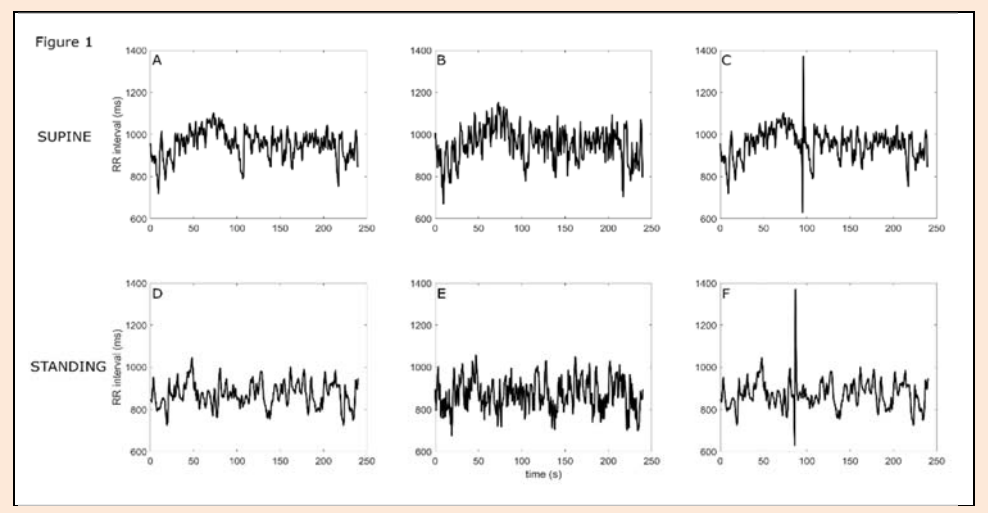


Figure 1. Panel A: raw RR-interval trace in the supine position with no artifact; Panel B: same trace as Panel A with each RR modified up or down by 50 ms, randomly; Panel C same trace as Panel A with one added artifact; Panels D, E and F identical to panels A, B and C, but in the standing position.

The random addition or subtraction to the successive RR intervals was repeated 5000 times for each case (i.e., 10, 30, 50 and 100 ms) in each of the 51 RR recording, resulting in 2,040,000 combinations.

The added artifact was a short-long sequence (i.e., an abnormally short RR immediately followed by an abnormally long RR in the time series), which is typically associated to spontaneous extra systole. The short RR was defined as the minimum RR value found in the series divided by 2 and the long RR was defined as the maximum RR value found in the series multiplied by 2. This short-long sequence was inserted in every possible position, that was first and second RR edited, then second and third etc. until the before-last and last RRs were edited, in the supine and the standing positions.

Figure 1B shows the same trace as in panel 1A with one of the 5000 iterations of the rand50 modifications. Figure 1C shows the same trace as in panel 1A with one example of artifact edition (out of the 252 possibilities in this trace).

Figure 1D shows a recording from the same participant as in Figure 1A, but in the standing position, again visibly containing no artifact. Data was edited similarly to that of the supine position. Figure 1E shows the same trace as in panel 1D with the rand50 modifications. Figure 1F shows the same trace as in panel 1E with the added artifact.

Automatic artifact correction

The process for artifact detection and correction is based on methods commonly found in the literature: abnormal heartbeats were corrected using cubic spline interpolation based on the normal heartbeats around (Lipponen and Tarvainen, 2019). An experienced researcher checked all corrections visually. The goal of this work was not to propose a new method for artifact detection and correction but to assess which HRV parameters are the most sensitive to artifacts and how many unresolved artifacts would significantly alter HRV interpretation.

Dataset 2 analysis

The aim of this analysis on dataset 2 was to determine how much artifacts, naturally present in RR recordings, would modify the HRV parameters. Therefore, the difference in each HRV parameter between the raw and the corrected trace was plotted versus the portion of the signal that had been corrected (cf. Figure 2 for supine and Figure 3 for standing position). For example, a recording of 250 RRs, that contained 10 artifacts had a corrected portion of 10/250 x 100 = 4%. For example, correcting those 10 artifacts in the RR recording changed RMSSD by 90 ms, then the point representing this particular recording will be plotted at [x,y] coordinates of [4,90]. This computation was repeated for the 178 recordings.

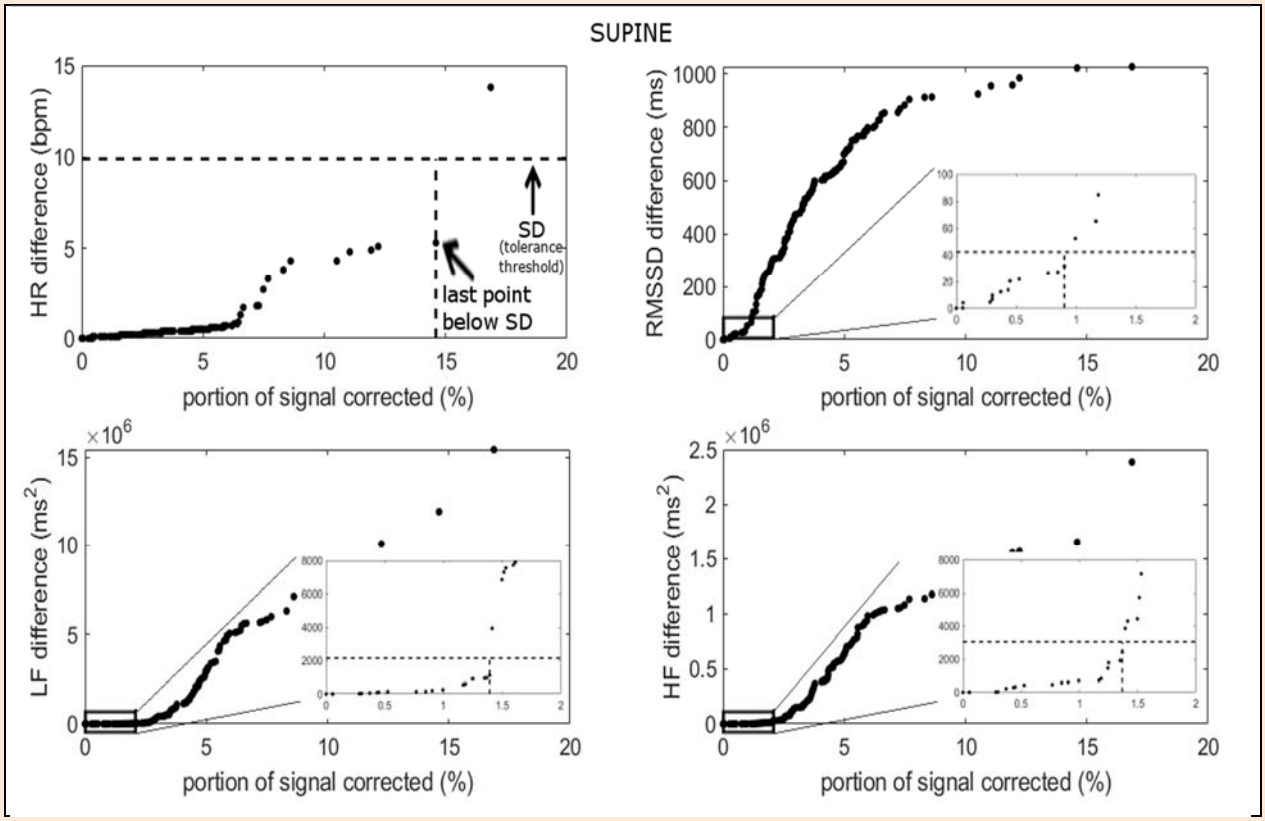


Figure 2. Supine position; x-axis: portion of the signal that contained artifacts before correction; y-axis: HR, RMSSD, LF or HF differences between raw trace and corrected trace. Horizontal dashed line is the standard deviation of the population, vertical dashed line crosses the last data point that is below the horizontal line, which is the least acceptable limit before artifacts changed the results more than the standard deviation of the population.

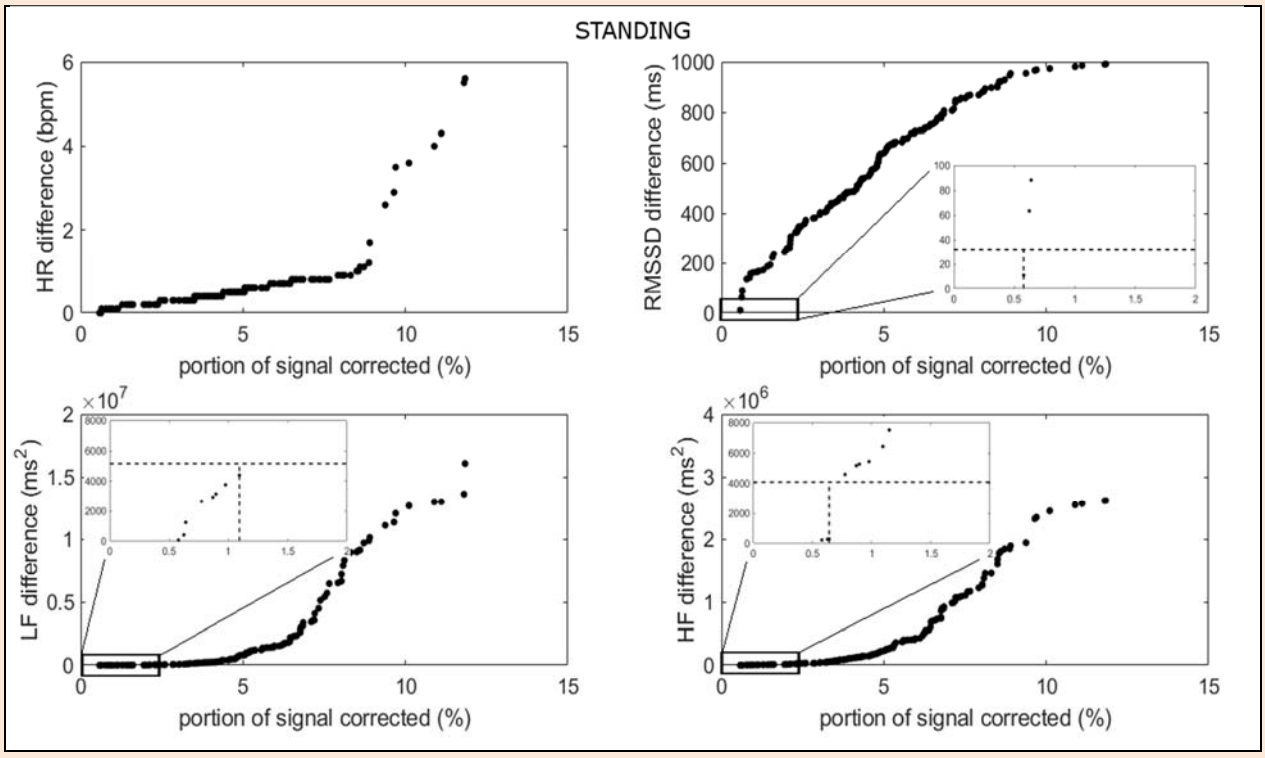


Figure 3. Standing position; x-axis: portion of the signal that contained artifacts before correction; y-axis: HR, RMSSD, LF or HF differences between raw trace and corrected trace. Horizontal dashed line is the standard deviation of the population, vertical dashed line crosses the last data point that is below the horizontal line, which is the least acceptable limit before artifacts changed the results more than the standard deviation of the population.

Each of the HRV parameters has a different dimension. It is therefore not easy to define a threshold common to all of them, above which the change in one parameter significantly alters interpretation. For example, a threshold of 5% difference in RMSSD between raw and corrected trace may be considered significant, whilst a 5% variation in HF power is largely insignificant. The horizontal dashed lines on Figure 2 and Figure 3 is the SD of our population and it is scaled to the dispersion of each parameter. Therefore, it was used to define the tolerance threshold. Below this tolerance threshold, the difference between raw and corrected HRV parameters is considered non-significant (< SD), whilst above this tolerance threshold the difference between raw and corrected HRV parameters is considered significant (> SD). Naturally, the greater the numbers of artifacts, the greater the portion of the signal corrected (x-axis of Figure 2 and Figure 3), the greater the difference in HRV parameters between raw and corrected trace (y-axis of Figure 2 and Figure 3), hence the monotonously increasing plots in Figure 2 and Figure 3. The vertical dashed lines on those figures denote the x-position of the last recording below the tolerance threshold and therefore the maximum portion of signal (%) that was corrected with a non-significant (< SD) variation in HRV parameters. The SD of the 178 recordings, therefore defining the tolerance thresholds were: HR: 10 and 12 bpm, RMSSD: 43 and 32 ms, SDNN: 47 and 47 ms, LF 2200 and 5200 ms^2, HF 3100 and 4100 ms^2 for supine and standing positions, respectively.

For clarity, inserted graphs within larger graphs are zooms on the portion of the graphs where the traces cross the tolerance thresholds (horizontal dashed lines in Figure 2 and Figure 3).

Statistical analysis

In dataset original trace, rand10, rand30, rand50, rand100 and artifact were compared using repeated measure one-way ANOVA with a significant level set at $p = 0.05$.

Results

Table 1 shows the selected HRV parameters of dataset 1 with the edited RR intervals. For rand10, rand30 and rand50, modifications were negligible for all HRV parameters. For rand100, LF and Tot were modified to a larger (and significant) extent than in rand10, rand30 and rand50. In the artifact condition, RMSSD and SDNN were significantly modified whilst this was not the case for LF, HF and Tot. All these modifications apply to both the supine and the standing positions. The artifact changed the results in the time domain much more than in the frequency domain.

Figure 2 shows the analysis of the 178 recordings of dataset 2 in the supine position. HR remained below the tolerance threshold (10 bpm) as long as artifacts were less than 14.5% of the total signal. Contradictory, LF, HF and SDNN remained below the tolerance threshold (2200, 3100 ms^2 and 47 ms for LF, HF and SDNN, respectively) until 1.4% of artifacts in the signal. RMSSD remained below the tolerance threshold (43 ms) until 0.9% of artifacts in the signal. In the supine position, RMSSD was more sensitive to artifact (0.9%) than SDNN, LF and HF

Table 1. HRV changes when editing RR intervals from a set of 51 orthostatic tests. Percentages compared to the original artifact-free recordings. Values are mean ± SD [min max].

		RMSSD, %	SDNN, %	LF, %	HF, %	Tot, %
supine	**rand10**	0 ± 0 [-1 1]	0 ± 0 [0 1]	-1 ± 0 [-2 1]	0 ± 1 [-3 4]	0 ± 0 [-1 1]
	rand30	0 ± 0 [-1 2]	0 ± 0 [-1 1]	0 ± 1 [-3 4]	1 ± 1 [-4 7]	0 ± 1 [-3 3]
	rand50	0 ± 0 [-1 2]	-1 ± 0 [-2 0]	-1 ± 1 [-6 4]	2 ± 1 [-2 6]	-1 ± 1 [-5 2]
	rand100	1 ± 0 [-1 3]	0 ± 0 [-1 2]	6 ± 12 [-3 337]*	-1 ± 2 [-8 14]	5 ± 18 [-4 668]*
	artifact	413 ± 2 [403 422]*	54 ± 0 [54 54]*	0 ± 1 [-1 7]	-1 ± 1 [-3 3]	+1 ± 1 [0 4]
standing	**rand10**	0 ± 0 [-1 1]	0 ± 0 [0 1]	1 ± 0 [-1 2]	-1 ± 1 [-4 3]	0 ± 0 [-1 2]
	rand30	0 ± 0 [-2 1]	0 ± 0 [-1 1]	-1 ± 1 [-4 3]	-1 ± 1 [-6 3]	-1 ± 1 [-3 2]
	rand50	0 ± 0 [-1 2]	0 ± 1 [-1 1]	2 ± 1 [-3 4]	2 ± 1 [-3 6]	1 ± 1 [-2 5]
	rand100	0 ± 0 [-1 2]	0 ± 0 [-1 1]	4 ± 13 [-5 441]*	2 ± 2 [-4 28]	5 ± 21 [-4 562]*
	artifact	269 ± 1 [264 272]*	47 ± 0 [47 47]*	1 ± 1 [0 8]	-1 ± 0 [-2 2]	0 ± 1 [0 9]

rand10 rand30 rand50 and rand100: randomly adding or subtracting 10, 30, 50 and 100ms to each RR interval; artifact: systematically adding a short-long sequence in each position possible in the RR times series. * $p < 0.05$ compared to the raw artifact-free trace.

(1.4%), which means that a smaller proportion of artifacts in the RR series was necessary for RMSSD to overcome the tolerance threshold.

Figure 3 shows the analysis of the 178 recordings of dataset 2 in the standing position. The difference in HR between raw and corrected trace was never higher than the tolerance threshold. Up to 1.1% and 0.6% of artifacts in the signal, LF and HF remained below the tolerance threshold (5200 and 4100 ms^2 for LF and HF, respectively). Up to 0.8% of artifacts in the signal, SDNN remained below the tolerance threshold (47 ms), whilst up to 0.6% of artifacts in the signal RMSSD remained below the tolerance threshold (32 ms). In the standing position, RMSSD and HF were more sensitive to artifact (0.6%) than SDNN (0.8%), and LF (1.1%); i.e., a smaller proportion of artifacts was necessary for RMSSD and HF to overcome the tolerance threshold, compared to other parameters.

Discussion

The main results of the present study were: 1) a single artifact affected RMSSD and SDNN to a larger extent than LF and HF; (2) modification of RR data points by 30 ms or less had negligible influence on RMSSD, SDNN, LF and HF; and 3) RMSSD was modified (i.e., distorted by >1SD) when 0.9% of the signal contained artifacts whilst this tolerance threshold increased to 1.4% for SDNN, LF and HF, in the supine position. Similarly, RMSSD and HF were affected when 0.6% of the signal contained artifacts whilst this threshold raised to 1.1% for LF, in the standing position.

Therefore, RMSSD seems more sensitive to the presence of artifacts than LF and HF. Unedited artifacts result in an increase in the randomness of short-term RR interval dynamics (Peltola et al., 2004), therefore affecting RMSSD more than other parameters as it is an index of short-term dynamics (Task Force, 1996). HF, which is also an index of short-term effect seems less sensitive than RMSSD at least in the supine position, because an isolated artifact does not alter the oscillation content of the RR. It takes several artifacts to sufficiently alter the general oscillation and modify the outcome of the frequency domain computations. This is in accordance with previous publications, where the effects of artifacts were clearly apparent even in simple measures of variance such as the standard deviation. Autoregressive modelling and frequency domain analysis can at least partly exclude aperiodic influences and hence may be less sensitive to occasional artifacts (Berntson and Stowell, 1998). In the present study, the signal should encompass at least 1.4% of artifacts to induce significant changes in LF and HF (i.e., greater than the SD of the present population). This corresponds to 3 artifacts for on average 213 heart beats in supine position and 4 artifacts for on average 275 heart beats in standing position (corresponding to 4 min of recording in each position). RMSSD, which is the most common parameter used by clinicians, is less satisfactory since it is a measure of spread and not a direct measurement of the deviation (Manis et al., 2005).

Regarding the artifact correction, the present work shows that modifying the RR intervals by 10 or 30 ms each (randomly up or down) does not alter the results of the HRV computation. In other words, the spread of RR intervals or the oscillations are not fundamentally altered. Therefore, automatic correction should focus on identifying and correcting single artifacts that would severely modify RMSSD rather than leaving unresolved artifacts. In this process, some normal RR intervals may be corrected, but as long as it is by less than 30 ms, the outcomes of HRV should not be altered. Many studies emphasize the importance of the artifact correction and appropriate editing for reliable HRV analyses. It would be important to standardize the editing practices within and between the studies (Tarkiainen et al., 2007). More comparative studies on large numbers of recordings are needed to define gold-standard recommendations for the suitable pre-processing and editing methods and for determining the maximum number of edited RR intervals in any short and long-term HRV analyses (Peltola, 2012).

RMSSD and SDNN are in ms, LF and HF are in ms^2; moreover, they have different reproducibility between and within study participants. Therefore, determining tolerance threshold is not easy and may depend on the type of application or population. Typically, a 5% change has certainly a different significance if it refers to RMSSD, LF or HF variations. Therefore, we decided to adopt the SD of our population as a threshold since it is representative of the dispersion within this group, independently of its clinical significance. However, those thresholds remain specific to the present dataset and more studies are needed to determine appropriately the suitable thresholds. HRV specialists would typically pick-up thresholds of few

milliseconds for RMSSD and few hundreds of ms^2 for LF and HF (Schmitt et al., 2015a). However, as the data on figures 2 and Figure 3 are monotonously increasing, picking different thresholds (as long as they are specific and scaled to each HRV parameter) would still result in RMSSD being more sensitive to artifacts than LF and HF.

In the present study, the analyzed windows were rather long (4-min each) whilst the recent literature focused on RMSSD computed from recordings as short as 60 s (Plews et al., 2012). Among other reasons, RMSSD is believed to be more robust than LF and HF and short recordings are more comfortable and less time-consuming for the users. An isolated artifact on a 4-min window alters RMSSD by 413% (table 1), a single artifact on a 60-s window induces even a bigger bias. However, it is four times less likely to occur than on a 4-min window. With a good, automated artifact correction (i.e., rather focused on over-correction than leaving unresolved artifacts), reporting time- and frequency-domain parameters in a comprehensive way should make HRV interpretation reliable and consistent.

The bias introduced by a given artifact may depend on its position in the RR time series (i.e., next to a local maximum, in a decreasing or increasing part of a waveform etc.). This has been documented elsewhere (Berntson and Stowell, 1998) in the literature and is beyond the scope of this article. Nevertheless, all artifact positions in each RR time series have been tested in the present work to avoid any bias that may have come from randomly selected positions.

Beyond the present considerations about the sensitivity of time- and frequency-domain parameters to artifacts, the clinical interests of combining RMSSD and LF-HF analyses have been demonstrated in previous publications, especially regarding fatigue type identification (Schmitt et al., 2015a) and HRV-guided training (Kiviniemi et al., 2007; Schmitt et al., 2018). Accurate HRV monitoring is essential in athletes and thus should rely both on time- and frequency-domain parameters.

Also, the HF band is related to the respiratory sinus arrhythmia and holds information that can hardly be seen on RMSSD only. Time-frequency analysis could represent an alternative for the assessment of cardiovagal regulation indexed by respiratory sinus arrhythmia (Mestanik et al., 2019). Finally, alternative techniques (i.e., not based on the debated LF-HF parameters), for the identification of the parasympathetic and sympathetic branches activity are increasingly proposed in the literature (Adjei et al., 2019; Rogers et al., 2021), but remain to be validated in athletes' follow-up.

Conclusion

In the literature, RMSSD was generally believed more robust and reliable than LF and HF for HRV analyses. However, the present work shows that it is more sensitive to artifacts than LF and HF. Time-domain parameters are very sensitive to a single artifact whilst frequency-domain parameters are less affected. Numerous little changes (< 30 ms in each RR) did not fundamentally change time- or frequency-domain parameters. Automatic correction systems should focus on slight over-correction of RRs rather than leaving unresolved artifacts. Beyond essential artifact correction, combining time- and frequency-domain analyses appears the wiser, safer and clinically relevant way to use HRV.

Acknowledgements

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. NB is a part time employee of be care SA. SY, JMV and GPM are unpaid members of the scientific committee of be.care SA. This work involved human participants which provided informed consent prior to participation in this study. This work was conducted according to the Declaration of Helsinki.

References

Adjei, T., von Rosenberg, W., Nakamura, T., Chanwimalueang, T. and Mandic, D.P. (2019) The ClassA Framework: HRV Based Assessment of SNS and PNS Dynamics Without LF-HF Controversies. *Frontiers in Physiology* **10**, 505. https://doi.org/10.3389/fphys.2019.00505.

Berntson, G.G., Quigley, K.S., Jang, J.F. and Boysen, S.T. (1990) An approach to artifact identification: application to heart period data. *Psychophysiology* **27(5)**, 586-598. https://doi.org/10.1111/j.1469-8986.1990.tb01982.x

Berntson, G.G. and Stowell, J.R. (1998) ECG artifacts and heart period variability: don't miss a beat!. *Psychophysiology* **35(1)**, 127-132. https://doi.org/10.1111/1469-8986.3510127

Bourdillon, N., Schmitt, L., Yazdani, S., Vesin, J. and Millet, G.P. (2017) Minimal Window Duration for Accurate HRV Recording in Athletes. *Frontiers Neuroscience* **11**, 456. https://doi.org/10.3389/fnins.2017.00456

Kiviniemi, A.M., Hautala, A.J., Kinnunen, H. and Tulppo, M.P. (2007) Endurance training guided individually by daily heart rate variability measurements. *European Journal of Applied Physiology* **101(6)**, 743-751. https://doi.org/10.1007/s00421-007-0552-2.

Le Meur, Y., Pichon, A., Schaal, K., Schmitt, L., Louis, J., Gueneron, J., et al. (2013) Evidence of parasympathetic hyperactivity in functionally overreached athletes. *Medicine and Science in Sports and Exercise* **45(11)**, 2061-2071. https://doi.org/10.1249/MSS.0b013e3182980125.

Lipponen, J.A. and Tarvainen, M.P. (2019) A robust algorithm for heart rate variability time series artefact correction using novel beat classification. *Journal of Medical Engineering & Technology* **43(3)**, 173-181. https://doi.org/10.1080/03091902.2019.1640306.

Manis, G., Alexandridi, A., Nikolopoulos, S. and Davos, K. (2005) The Effect of White Noise and False Peak Detection on HRV Analysis. in Proceedings of the 1st International Workshop on Biosignal Processing and Classification. The First International Workshop on Biosignal Processing and Classification, Barcelona, Spain: *SciTePress - Science and and Technology Publications*, 161-166. https://doi.org/10.5220/0001195301610166.

Mestanik, M., Mestanikova, A., Langer, P., Grendar, M., Jurko, A., Sekaninova, N., et al. (2019) Respiratory sinus arrhythmia - testing the method of choice for evaluation of cardiovagal regulation. *Respiratory Physiology & Neurobiology* **259**, 86-92. https://doi.org/10.1016/j.resp.2018.08.002.

Peltola, M.A. (2012) Role of editing of R-R intervals in the analysis of heart rate variability. *Frontiers in Physiology* **3**, 148. https://doi.org/10.3389/fphys.2012.00148.

Peltola, M.A., Seppänen, T., Mäkikallio, T.H. and Huikuri, H.V. (2004) Effects and significance of premature beats on fractal correlation properties of R-R interval dynamics. *Annals of Noninvasive Electrocardiology* **9(2)**, 127-135. https://doi.org/10.1111/j.1542-474X.2004.92531.x.

Pitzalis, M.V., Mastropasqua, F., Massari, F., Forleo, C., Di Maggio, M., Passantino, A., et al. (1996) Short- and long-term reproducibility of time and frequency domain heart rate variability measurements in normal subjects. *Cardiovascular Research* **32(2)**, 226-233. https://doi.org/10.1016/0008-6363(96)00086-7.

Plews, D.J., Laursen, P.B., Kilding, A.E. and Buchheit, M. (2012) Heart rate variability in elite triathletes, is variation in variability the key to effective training? A case comparison. *European Journal of Applied Physiology* **112(11)**, 3729-3741. https://doi.org/10.1007/s00421-012-2354-4.

Rincon Soler, A.I., Silva, L.E.V., Fazan, R. and Murta, L.O. (2018) The impact of artifact correction methods of RR series on heart rate variability parameters. *Journal of Applied Physiology* **124(3)**, 646-652. https://doi.org/10.1152/japplphysiol.00927.2016.

Rogers, B., Mourot, L., Doucende, G. and Gronwald, T. (2021) Fractal correlation properties of heart rate variability as a biomarker of endurance exercise fatigue in ultramarathon runners. *Physiological Reports* **9(14)**, e14956. https://doi.org/10.14814/phy2.14956.

Saboul, D. and Hautier, C. (2019) A New Algorithm to Reduce and Individualize HRV Recording Time. *Journal of Medical Systems* **43(3)**, 45. https://doi.org/10.1007/s10916-019-1167-y.

Salo, M.A., Huikuri, H.V. and Seppänen, T. (2001) Ectopic beats in heart rate variability analysis: effects of editing on time and frequency domain measures. *Annals of Noninvasive Electrocardiology* **6(1)**, 5-17. https://doi.org/10.1111/j.1542-474X.2001.tb00080.x

Sandercock, G.R.H., Bromley, P. and Brodie, D.A. (2004) Reliability of three commercially available heart rate variability instruments using short-term (5-min) recordings. *Clinical Physiology and Functional Imaging* **24(6)**, 359-367. https://doi.org/10.1111/j.1475-097X.2004.00584.x.

Schmitt, L., Bouthiaux, S. and Millet, G.P. (2021) Eleven Years' Monitoring of the World's Most Successful Male Biathlete of the Last Decade. *International Journal of Sports Physiology and Performance* **16(6)**, 900-905. https://doi.org/10.1123/ijspp.2020-0148.

Schmitt, L., Regnard, J., Desmarets, M., Mauny, F., Mourot, L., Fouillot, J.-P., et al. (2013) Fatigue shifts and scatters heart rate variability in elite endurance athletes. *Plos One* **8(8)**, e71588. https://doi.org/10.1371/journal.pone.0071588.

Schmitt, L., Regnard, J. and Millet, G.P. (2015a) Typology of "Fatigue" by Heart Rate Variability Analysis in Elite Nordic-skiers. *International Journal of Sports Medicine* **36(12)**, 999-1007. https://doi.org/10.1055/s-0035-1548885.

Schmitt, L., Regnard, J., Parmentier, A.L., Mauny, F., Mourot, L., Coulmy, N., et al. (2015b) Monitoring Fatigue Status with HRV Measures in Elite Athletes: An Avenue Beyond RMSSD?. *Frontiers in Physiology* **6**, 343. https://doi.org/10.3389/fphys.2015.00343.

Schmitt, L., Willis, S.J., Fardel, A., Coulmy, N. and Millet, G.P. (2018) Live high-train low guided by daily heart rate variability in elite Nordic-skiers. *European Journal of Applied Physiology* **118(2)**, 419-428. https://doi.org/10.1007/s00421-017-3784-9.

Stapelberg, N.J.C., Neumann, D.L., Shum, D.H.K., McConnell, H. and Hamilton-Craig, I. (2018) The sensitivity of 38 heart rate variability measures to the addition of artifact in human and artificial 24-hr cardiac recordings. *Annals of Noninvasive Electrocardiology* **23(1)**. https://doi.org/10.1111/anec.12483

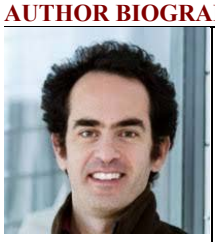
Tarkiainen, T.H., Kuusela, T.A., Tahvanainen, K.U.O., Hartikainen, J.E.K., Tiittanen, P., Timonen, K.L., et al. (2007) Comparison of methods for editing of ectopic beats in measurements of short-term non-linear heart rate dynamics. *Clinical Physiology and Functional Imaging* **27(2)**, 126-133. https://doi.org/10.1111/j.1475-097X.2007.00726.x.

Task Force (1996) Heart rate variability. Standards of measurement, physiological interpretation, and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. *European Heart Journal* **17(3)**, 354-381. https://pubmed.ncbi.nlm.nih.gov/8598068/

Key points

- RMSSD and SDNN are more sensitive to a single artifact than LF and HF
- Modification of RR data points by 30 ms or less had negligible influence on RMSSD, SDNN, LF and HF
- Automatic correction systems should focus on slight over-correction of RRs rather than leaving unresolved artifacts
- Combining time- and frequency-domain analyses appears the wiser, safer and clinically most relevant way to use HRV

AUTHOR BIOGRAPHY



Nicolas BOURDILLON
Employment
Research assistant at UNIL and Chief Research Officer at be.care S.A.
Degree
PhD
Research interests
Cardiovascular physiology, fatigue, altitude, signal processing
E-mail: nicolas.bourdillon@unil.ch

Sasan YAZDANI
Employment
Senior Bio-medical Signal Processing Manager, Aktiia
Degree
PhD
Research interests
Biomedical signal processing
E-mail: sassan.yazdani@gmail.com

Jean-Marc VESIN
Employment
Senior scientist, EPFL
Degree
PhD
Research interests
Biomedical signal processing, ECG, PPG, atrial fibrillation, nonlinear signal modeling and prediction
E-mail: jean-marc.vesin@epfl.ch

Laurent SCHMITT
Employment
Head of Research-Expertise-Performance department at the French national cross-country skiing center
Degree
PhD
Research interests
Performance, fatigue
E-mail:
laurent.schmitt@ensm.sports.gouv.fr

Grégoire P. MILLET
Employment
Associate Professor, Deputy Head of Service
Degree
PhD
Research interests
Fatigue, hypoxia, training, performance
E-mail: gregoire.millet@unil.ch

✉ **Nicolas Bourdillon**
ISSUL, Institute of Sport Sciences, University of Lausanne, Lausanne, Switzerland