SCIENTIFIC REPORTS

OPEN

The application of Lempel-Ziv and Titchener complexity analysis for equine telemetric electrocardiographic recordings

Vadim Alexeenko[1,2], James A. Fraser[2], Alexey Dolgoborodov[3], Mark Bowen[4], Christopher L.-H. Huang[2,6], Celia M. Marr[5] & Kamalan Jeevaratnam[1,2]

The analysis of equine electrocardiographic (ECG) recordings is complicated by the absence of agreed abnormality classification criteria. We explore the applicability of several complexity analysis methods for characterization of non-linear aspects of electrocardiographic recordings. We here show that complexity estimates provided by Lempel-Ziv '76, Titchener's T-complexity and Lempel-Ziv '78 analysis of ECG recordings of healthy Thoroughbred horses are highly dependent on the duration of analysed ECG fragments and the heart rate. The results provide a methodological basis and a feasible reference point for the complexity analysis of equine telemetric ECG recordings that might be applied to automate detection of equine arrhythmias in equine clinical practice.

It has been widely recognised that cardiac arrhythmias pose a significant health risk for athletes, both human[1] and equine[2,3]. They could not only impair performance but might go on to trigger potentially fatal cardiac episodes. Equine and human athletes exhibit parallel lifestyle exercise patterns taking the form of intense training which reaches its peak during races or other competitive performances and is followed by substantial reduction in intensity of exercise during subsequent lifetime retirement period. Both attain similar heart rates during individual training episodes apart from the wider equine than human heart rate variations (20 to 220 bpm).

Cardiac arrhythmias result from a diverse set of irregularities in impulse initiation and conduction or combination of both. In healthy subjects the electrical signal initiation is confined to well-defined cardiac regions such as sinoatrial or atrioventricular node, abnormalities in which result in well-defined electrocardiographic (ECG) changes. The resulting alterations in observed cardiac electrical activity may appear as directly identifiable changes in the ECG waveform both in horses and humans, particularly in the form of alterations in sinus rate, as the consequences of major anatomical disruptions in conducting or contractile tissue. More subtle ECG changes could also arise at the tissue level from atrial and/or ventricular myocardial diseases including channelopathies, alterations in calcium homeostasis and tissue heterogeneities[4]. Early detection of such changes might precede the first arrhythmic episode and therefore act as its predictor. The wide-spread use of computers and digital ECG recordings has facilitated the development of the computer-based methods of arrhythmia detection in human subjects[5–7].

Although ECG-based screening programmes appropriate for the detection of risk of sudden cardiac deaths are available for human athletes[8] such screening programmes do not exist for equine athletes. This in part reflects the lack of agreed classification criteria for abnormality, particularly in detection of premature complexes. Thus, subjective judgment of what is considered normal or abnormal affects the reliability of equine ECG-based diagnostics. To eliminate this subjectivity various objective measures have been proposed. In our previous paper we explored use of ECG restitution analysis to evaluate the function of equine cardiovascular system[9]. The restitution analysis has been shown to indicate the integrated mechanoelectrical impacts on the heart and of arrhythmia

[1]Faculty of Health and Medical Sciences, University of Surrey, Guildford, GU2 7AL, United Kingdom. [2]Physiological Laboratory, University of Cambridge, Cambridge, CB2 3DY, United Kingdom. [3]Seven Bridge Genomics, Boston, MA, 02142, USA. [4]Faculty of Medicine & Health Sciences, University of Nottingham, Nottingham, NG7 2UH, United Kingdom. [5]Rossdales Equine Hospital and Diagnostic Centre, Exning, CB8 7NN, Suffolk, United Kingdom. [6]Division of Cardiovascular Biology, Department of Biochemistry, University of Cambridge, Cambridge, CB2 1QW, United Kingdom. Correspondence and requests for materials should be addressed to K.J. (email: drkamalanjeeva@gmail.com)

liability[10]. This method requires the recording of good quality ECG at high heart rates during the exercise, when electrical signal from the heart might be obscured by electrical activity of other muscles.

A more direct approach to quantify the anomalies in generation and propagation of electrical signal in the heart is to use signal complexity estimation techniques. These techniques are shown to be a sensitive tool to estimate the irregularity of various bioelectrical signals, including neuronal spiking and electroencephalographic (EEG) records[11–14]. The inherently chaotic nature of the heartbeat (as pointed out by Goldberger in 1991)[15] makes complexity analysis a suitable tool to assess its stochasticity. Various complexity estimators were employed to identify arrhythmia[16], to analyse circadian variability of heart rate variation[17,18], for characterization of intermittent complexity variations during ventricular fibrillation[19] as well as for prediction of the onset of atrial fibrillation episode[20]. Complexity estimation techniques have a benefit of not requiring high heart rate in analysed data.

These techniques are based on the seminal work of Kolmogorov[21] who suggested that the disorderliness (complexity) of a string of a symbols might be described as the length of the shortest possible computer program capable of generating such string. Lempel and Ziv later demonstrated that complexity can be feasibly linked to the gradual build-up of new patterns along the given sequence[22]. Subsequently, a mathematical method was developed to apply the concept of Shannon entropy[23] to repetitive time series, resulting in development of approximate entropy[24] and sample entropy[25] metrics, as well as some other variations, such as wavelet entropy[26]; see the recent paper of Dharmaprani *et al.*[27] for a review. All these metrics are based on an analysis of repetitive patterns in the time series and have been used for analysis of physiological signals. In this study we primarily focus on analysis of metrics based on Lempel-Ziv string compression analyses, as it has been demonstrated that they have several-fold lower computational complexity compared to approximate entropy or sample entropy[28] and also have no arbitrary tuning parameters which might affect their performance[29].

The key idea of the initial Lempel and Ziv's method (usually denoted as LZ 76) is to decompose the source string of symbols to a set of such patterns ('terms' or 'factors') which are sufficient to rebuild the source string by a machine performing copy and insertion operations[30]. The number of these patterns obtained in the decomposition step (LZ vocabulary size) is directly proportional to the complexity of the source string, and therefore to the entropy of its source. Detailed analysis of LZ decomposition method and its relation to Kolmogorov's complexity can be found in the published literature[14,22,30–32].

To obtain meaningful estimates of complexity values using these algorithms requires simultaneously that a sufficient amount of data is analysed and that the analysed parameter should be stationary during the data collection period[33,34]. This may lead to severe practical limitations in the time resolution of analysis in a quasiperiodic signal like the ECG as the absolute minimal amount of data is estimated to be in the range from 10 heartbeats for LZ 76 up to 80 heartbeats for sample entropy[35]. Recording of 80 equine heartbeats at rest might require up to four minutes of recording time raising the possibility of heart rate alterations originating outside of the cardiovascular system[36].

A slightly modified and faster version of Lempel-Ziv decomposition was subsequently proposed to develop the Lempel-Ziv 1978 data compression algorithm[37]. This decomposition method (LZ 78) might be used as a complexity estimation tool as well. Both methods feature incomplete parsing of the final part of the analysed symbolic string, which introduces additional variability in the case of string shorter than 10000 symbols[33,38]. Although the error caused by this variability could be decreased by analysing longer data strings, such increase is often undesirable for analysis of bioelectrical signals as it limits its temporal resolution. An alternative way to decrease this error is to use a decomposition algorithm capable of parsing input data without any incomplete parts being left, for example one suggested by Titchener[39].

It should be noted that although complexity estimators do produce estimates of the regularity of a signal, there is no guarantee that a particular estimator will reliably process an arbitrary type of data input. Apart from above-mentioned difficulties with analysis of "short" data sets (typical for the ECG data obtained in primary care settings), certain classes of "long" ($\gg 10^5$ samples) data series[40] may also be incorrectly classified by these estimators. Such problems warrant in-depth study of the suitability of a given complexity estimator to process the data before the method can be used to investigate a relevant biological problem. Probably the most important factor to be addressed in such a study would be the behaviour of an estimator in range of signal durations easily collectable in clinical practice (less than 60 seconds) and the dependence of complexity values on the heart rate.

As complexity estimation algorithms deal with sequences of symbols (strings) rather than floating-point data obtained in physiological recordings, several preparatory steps are required before the ECG can be subjected to complexity analysis. These steps typically include low-pass or band-pass filtering, baseline wander correction and resampling followed by data granulation involving conversion of the sequence of numerical ECG voltage readings to a symbolic string. Careful selection of data conditioning parameters, including sampling rate and sample length are required to ensure reproducibility of the results. The present paper describes use and feasibility of complexity estimation methods for analysis of equine ECG. In doing so we emerged with a clear methodological approach and convenient option to use in development of risk prediction algorithms for equine cardiac rhythm abnormalities.

Results

Dependence of ECG complexity values on the ECG strip length.
To estimate the dependence of the computed complexity on the length of the analysed strip we selected a convenient sample of contiguous 120-sec ECG recordings excised from longer raw electrocardiograms of 15 individual horses. Recordings were selected to cover the range of heart rates from 32 to 145 bpm. Each 120-sec recording was then used to generate 11 sets of 60 random strips of pre-defined shorter length (20, 24, 28, 35, 42, 50, 60, 72, 86, 100 and 118 sec; total 660 strips per horse/900 strips for each strip length) which were subjected to complexity analysis. As demonstrated by a representative sample (Fig. 1a) we observed that resulting complexity values decreased in absolute value and variability with increase of the analysed strip length. To quantify the error in determination of these values we calculated
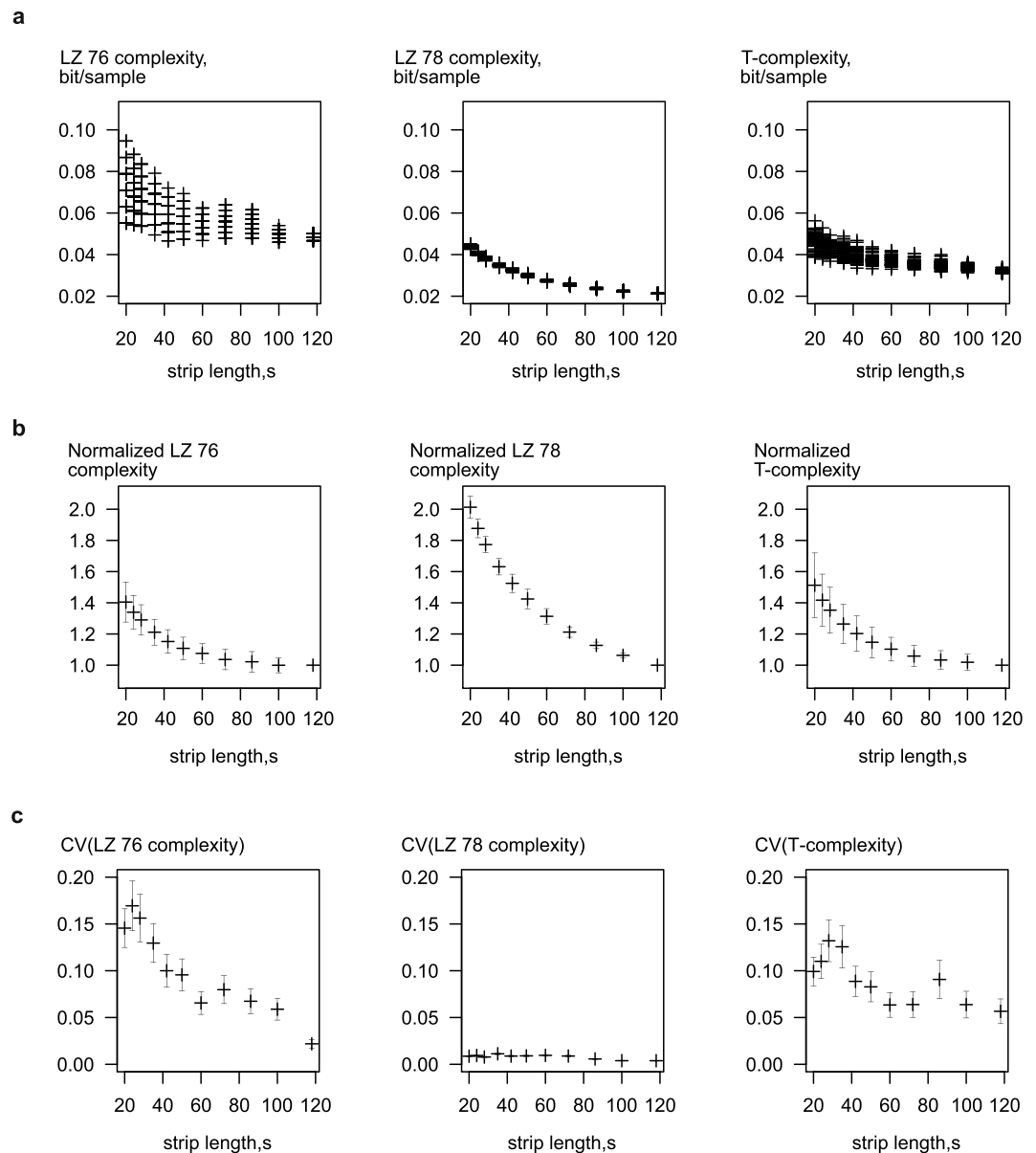
Figure 1. Effect of signal length on the complexity values returned by different estimators. (**a**) Distribution of complexity values obtained by analysis of different length samples extracted from a 120-sec fragment of telemetric equine ECG recording. Sixty random samples from each of the 15 horses analysed for each strip length. (**b**) Dependence of complexity values normalized to complexity at the maximal strip length (n = 15). (**c**) Effect of sample length on the coefficient of variation of different complexity values (n = 15).

ratios of complexity values obtained for the shorter strips in relation to complexity of the longest (118-sec) strips, thus eliminating dependence on the baseline complexity variations in individual subjects. The maximum increase in the shortest (20 sec) strips was up to 1.5-2-fold compared to the longest (118 sec) strips, indicating that the analysis of very short ECG strips is prone to systematic overestimation of its complexity (Fig. 1b). Additionally we have found that variation coefficients (ratio of the standard deviation to the mean which shows the extent of variability in relation to the mean of the sample) in each 60-strip group were also markedly increased in strips shorter than 60 sec (Fig. 1c). Therefore 60-second strips were used for subsequent analyses.

Dependence of ECG complexity on heart rate.

The initial analysis of complexity dependence on the analysed strip length allowed us to evaluate the possible influence of heart rate on the ECG complexity. Due to the limitations of the data collection process for this study each subject was offering data only for a limited range of heart rates (Fig. 2a). Therefore, all available 60-sec strips extracted from all recordings were combined in one data set and grouped at 5 bpm heart rate intervals. The average LZ 76 and Titchener complexity values (but not LZ 78 values) have demonstrated the asymptotically increasing behaviour at high heart rates (Fig. 2b).
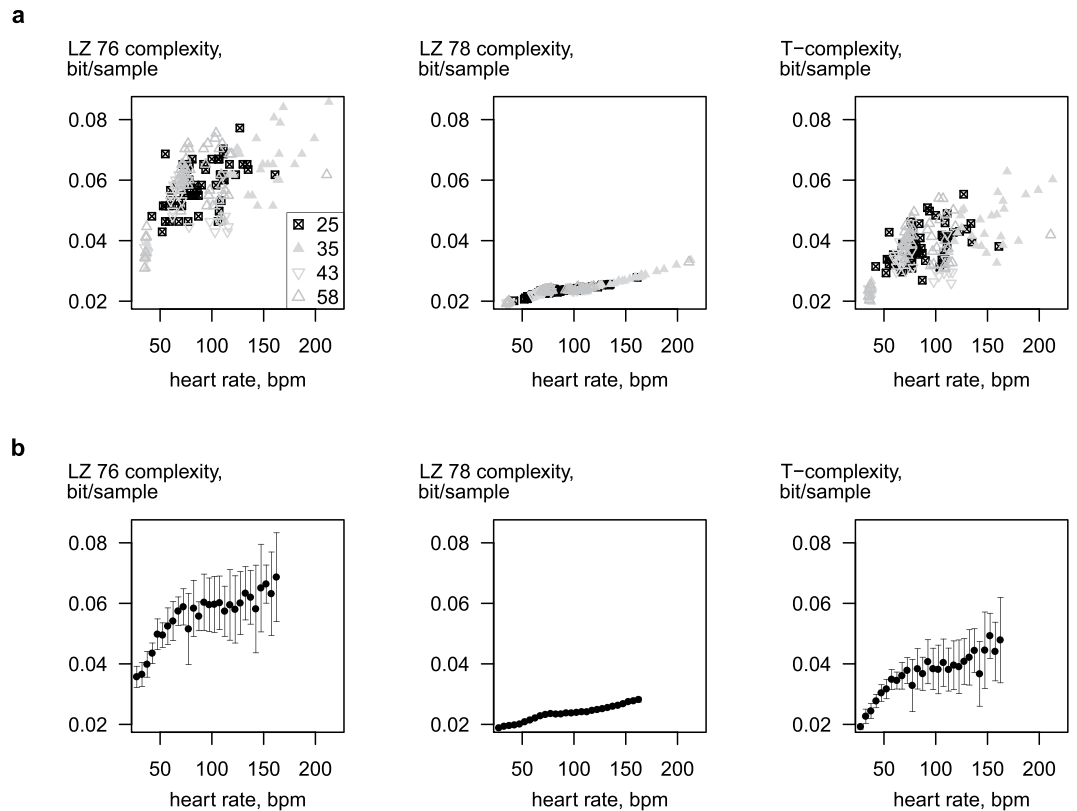
Figure 2. Dependence of ECG complexity on heart rate. (**a**) Representative data from four horses. (**b**) Average data from 51 horses grouped at 5 bpm heart rate ranges; only groups with three or more observations are shown.

Discussion

Our study has demonstrated that random short (60–100 sec) samples extracted from ambulatory equine ECG recordings may be used to evaluate ECG complexity. There is a clear relationship between length of the analysed strip and its complexity as well as between heart rate and complexity. All three complexity estimation methods show relatively small variabilities in their results at heart rates less than 70 bpm. At higher heart rates they exhibit a markedly different behaviour: LZ 76 complexity and T-complexity values exhibit increased spreads of complexity values with average complexity values reaching plateau. On the other hand, LZ 78 complexity values show remarkably little variation and very little increase at high heart rates. We might speculate that such different behaviour might be due to the different algorithms being sensitive to different physiological phenomena, possibly providing the foundation for the new biomarkers. The asymptotic behaviour of LZ 76 and T-complexity highlights the possibility of introduction of heart-rate corrected complexity metrics, which might be very useful to compare the data obtained at different heart rates. However, the additional theoretical and experimental studies which are required to introduce such correction formula were beyond the scope of this study. Such corrections might facilitate the direct comparison of complexity obtained at different heart rates and in different studies. The validity of such correction remains to be demonstrated in a future studies involving subjects with known pathologies associated with alterations in ECG complexity, primarily atrial fibrillation[17,41].

Our results broadly agree with the recent study published by Zhang et al.[42] in that ECG strips longer than 40 seconds are required to perform complexity analysis. In our study somewhat longer (60 sec) strips were required to stabilise complexity values. This might be due to the different coarse-graining technique we employ (beat detection instead of threshold crossing). Our choice of "beat detection" coarse-graining was primarily due to the high variability of the equine ECG and its absence of the horizontal T-P intervals which is typical for the human ECG and provides a convenient place to detect increased noise from atria. It could be expected that the high variability of the equine ECG might require development of appropriate techniques for baseline subtraction to facilitate threshold-crossing coarse-graining. Alternatively, other coarse-graining techniques might be developed for the same purpose: for example, one that relies on detection of P, Q and T wave boundaries. Such metrics would be less sensitive to noise than threshold-crossing methods and may provide more information on the excitation wave travel than beat-detection technique. Additional valuable information might possibly be supplied by emerging complexity analyses based on the second-moment statistics[43]. We expect that such easily automated analyses, probably in conjunction with currently used biomarkers might potentially be able to serve as sensitive tools for cardiovascular diagnostics. They would supplement, if not potentially replace the resource-consuming "fishing for irregular heartbeat"[44] which is currently used for diagnosis of paroxysmal atrial fibrillation and other hard-to-detect intermittent arrhythmias in horses.
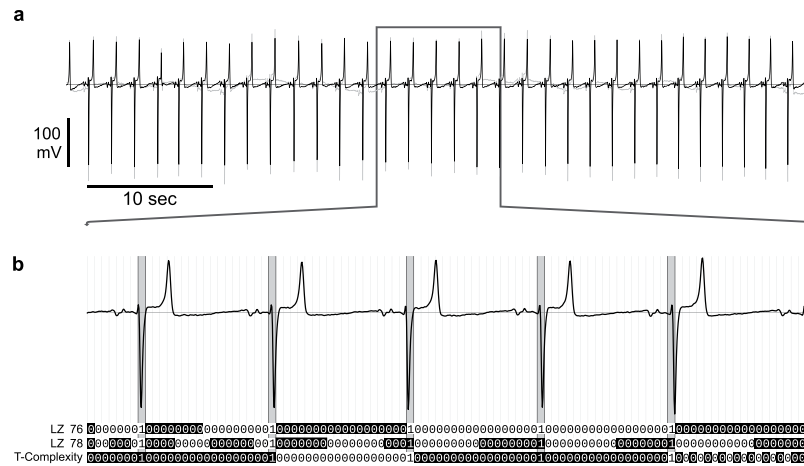
Figure 3. Equine ECG complexity analysis method. (**a**) Typical segment of equine ambulatory ECG. Grey, original signal, black – signal after resampling and filtration. (**b**) Conversion of ECG recording to a beat-detection binary string and subsequent complexity analyses. Beat detections shown by vertical grey bars. Decomposition of the resulting beat-detection binary string to the individual factors is indicated by alternating black/white text background. See text for the description of the decomposition methods.

Conclusion

The present *in vivo* study provides a methodological basis for the complexity analysis of equine telemetric ECG recordings.

Methods

Subject recruitment. Based on the ethical assessment review checklist by the Non-ASPA Sub-Committee at the University of Surrey, the study did not require an ethical review and received appropriate faculty level approval. Non-invasive ECG recordings were collected at Rossdales Equine Hospital and Diagnostic Centre (Newmarket, Suffolk, United Kingdom) during routine clinical work up from 51 Thoroughbred horses in race training. Data were anonymised prior to analysis. ECGs had been recorded while horses were exercised as part of their clinical investigations. All horses were of racing age at the time of testing. None showed clinically significant cardiac abnormalities.

Data recording. Each horse was atraumatically fitted with a telemetric ECG recorder (Televet 100, Engel Engineering Services GmbH, Germany). Data were recorded in continuous episodes lasting 10–50 minutes, before and during a period of acceleration from walk to canter. This emulated incremental pacing protocol has previously been applied in studies of cardiac function *in vitro*[45,46] and yielded ECG strips offering range of relatively steady incremental heart rates. The Televet 100 recorder has signal bandwidth of 0.05–100 Hz and sampling rate of 500 Hz.

Data preparation. The original data files were exported to CSV files with the TeleVet software. A custom software package written in GNU R, version 3.2.3[47] was used to split the long ECG recordings and extract artefact-free 60-sec strips of lead I ECG. These strips were consequently downsampled to 125 Hz and filtered using the digital band-pass Butterworth filter. The empirically selected cut-off frequencies of 0.13 Hz and 23 Hz allowed elimination of high-frequency noise and baseline drift (Fig. 3a). The filtered recordings were converted to binary strings using a custom beat detection algorithm implemented in R (Fig. 3b). The results of the beat-detection analysis were checked for the detection errors and strips with false detections or missed beats were excluded from the subsequent analysis.

Complexity analysis. Estimation of the binary string complexities was facilitated by a custom implementation of a complexity estimator developed in C++ for the Linux operating system, using a Qt framework (http://www.qt.io). The program simultaneously performed complexity analysis using three previously published methods: Lempel-Ziv '76[22], Lempel-Ziv '78[37] and Titchener T-complexity[39]. All three methods estimate the complexity of the symbolic string by identifying the number of sub-strings needed to build the source string (indicated by alternating black and white background in Fig. 3b). These methods differ by the algorithms of decomposition of the source strings to sub-strings; detailed descriptions of these algorithms can be found in corresponding publications. To eliminate the dependency of complexity values on the length of the source string (n), LZ 76 complexity values were normalised to the $n/\log_2(n)$ value[30] and LZ 78 values were normalised to sequence length. For the T-complexity, average entropy values were used.

Statistical analyses. Parametric data are expressed as mean ± standard deviation of mean. Statistical analyses and plotting were done using GNU R[47].

References

1. Sharma, S., Merghani, A. & Mont, L. Exercise and the heart: The good, the bad, and the ugly. *Eur. Heart J.* **36**, 1445–1453 (2015).
2. Slack, J., Boston, R. C., Soma, L. R. & Reef, V. B. Occurrence of cardiac arrhythmias in Standardbred racehorses. *Equine Vet. J.* **47**, 398–404 (2015).
3. Ryan, N., Marr, C. M. & McGladdery, A. J. Survey of cardiac arrhythmias during submaximal and maximal exercise in Thoroughbred racehorses. *Equine Vet. J.* **37**, 265–8 (2005).
4. Tse, G. Mechanisms of cardiac arrhythmias. *J.Arrhythm.* **32**, 75–81 (2016).
5. Goldberger, A. L. *et al.* PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. *Circulation* **101**, 215–220 (2000).
6. Clifford, G. *et al.* AF Classification from a Short Single Lead ECG Recording: the Physionet Computing in Cardiology Challenge 2017. **44**, 1–4 (2017).
7. Julián, M., Alcaraz, R. & Rieta, J. J. Comparative assessment of nonlinear metrics to quantify organization-related events in surface electrocardiograms of atrial fibrillation. *Comput. Biol. Med.* **48**, 66–76 (2014).
8. Emery, M. S. & Kovacs, R. J. Sudden Cardiac Death in Athletes. *JACC Hear. Fail.* **6**, 30–40 (2018).
9. Li, M. *et al.* Cardiac electrophysiological adaptations in the equine athlete-Restitution analysis of electrocardiographic features. *Plos One* **13**, e0194008 (2018).
10. Fossa, A. A. Beat-to-beat ECG restitution: A review and proposal for a new biomarker to assess cardiac stress and ventricular tachyarrhythmia vulnerability. *Ann. Noninvasive Electrocardiol.* **22**, 1–11 (2017).
11. Abasolo, D., James, C. J. & Hornero, R. Non-linear analysis of intracranial electroencephalogram recordings with approximate entropy and Lempel-Ziv complexity for epileptic seizure detection. *Conf. Proc. IEEE Eng Med. Biol. Soc.* **2007**, 1953–1956 (2007).
12. Jouny, C. C. & Bergey, G. K. Characterization of early partial seizure onset: frequency, complexity and entropy. *Clin. Neurophysiol.* **123**, 658–669 (2012).
13. Hornero, R., Abasolo, D., Escudero, J. & Gomez, C. Nonlinear analysis of electroencephalogram and magnetoencephalogram recordings in patients with Alzheimer's disease. *Philos. Trans. A Math. Phys. Eng Sci.* **367**, 317–336 (2009).
14. Amigo, J. M., Szczepanski, J., Wajnryb, E. & Sanchez-Vives, M. V. Estimating the entropy rate of spike trains via Lempel-Ziv complexity. *Neural Comput.* **16**, 717–736 (2004).
15. Goldberger, A. Is the Normal Heartbeat Chaotic or Homeostatic? *Physiology* **6**, 87–91 (1991).
16. Udhayakumar, R. K., Karmakar, C., Li, P. & Palaniswami, M. Effect of embedding dimension on complexity measures in identifying Arrhythmia. *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS* **2016-Octob**, 6230–6233 (2016).
17. Cabiddu, R. *et al.* Are complexity metrics reliable in assessing HRV control in obese patients during sleep? *Plos One* **10**, 1–15 (2015).
18. Ferrario, M., Signorini, M. G. & Cerutti, S. Complexity analysis of 24 hours heart rate variability time series. *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2**, 3956–9 (2004).
19. Schlemmer, A., Baig, T., Luther, S. & Parlitz, U. Detection and characterization of intermittent complexity variations in cardiac arrhythmia. *Physiol. Meas.* **38**, 1561–1575 (2017).
20. Chesnokov, Y. V. Complexity and spectral analysis of the heart rate variability dynamics for distant prediction of paroxysmal atrial fibrillation with artificial intelligence methods. *Artif. Intell. Med.* **43**, 151–165 (2008).
21. Kolmogorov, A. N. Three approaches to the concept of the amount of information. *Probl. Inform. Transmission* **1**, 1–7 (1965).
22. Lempel, A. & Ziv, J. Complexity of Finite Sequences. *IEEE Trans. Inf. Theory* **22**, 75–81 (1976).
23. Shannon, C. E. A Mathematical Theory of Communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
24. Pincus, S. M. Approximate entropy as a measure of system complexity. *Proc. Natl. Acad. Sci. USA* **88**, 2297–301 (1991).
25. Richman, J. S. & Moorman, J. R. Physiological time-series analysis using approximate entropy and sample entropy. *Am. J. Physiol. Circ. Physiol.* **278**, H2039–H2049 (2000).
26. Alcaraz, R. & Rieta, J. J. Application of wavelet entropy to predict atrial fibrillation progression from the surface ECG. *Comput. Math. Methods Med.* **2012** (2012).
27. Dharmaprani, D. *et al.* Information theory and Atrial Fibrillation (AF): A review. *Front. Physiol.* **9**, 1–19 (2018).
28. Kedadouche, M., Thomas, M., Tahan, A. & Guilbault, R. Nonlinear Parameters for Monitoring Gear: Comparison between Lempel-Ziv, Approximate Entropy, and Sample Entropy Complexity. *Shock Vib.* **2015** (2015).
29. Yentes, J. M. *et al.* The appropriate use of approximate entropy and sample entropy with short data sets. *Ann. Biomed. Eng.* **41**, 349–365 (2013).
30. Kaspar, F. & Schuster, H. G. Easily calculable measure for the complexity of spatiotemporal patterns. *Phys. Rev. A* **36**, 842–848 (1987).
31. Aboy, M., Hornero, R., Abásolo, D. & Álvarez, D. Interpretation of the Lempel-Ziv complexity measure in the context of biomedical signal analysis. *IEEE Trans. Biomed. Eng.* **53**, 2282–2288 (2006).
32. Artan, N. S. EEG analysis via multiscale Lempel-Ziv complexity for seizure detection. *Conf. Proc…. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. IEEE Eng. Med. Biol. Soc. Annu. Conf.* **2016**, 4535–4538 (2016).
33. Speidel, U. A note on the estimation of string complexity for short strings. *ICICS 2009 - Conf. Proc. 7th Int. Conf. Information, Commun. Signal Process*, https://doi.org/10.1109/ICICS.2009.5397536 (2009).
34. Hu, J., Gao, J. & Principe, J. C. Analysis of biomedical signals by the Lempel-Ziv complexity: The effect of finite data size. *IEEE Trans. Biomed. Eng.* **53**, 2606–2609 (2006).
35. Balasubramanian, K. & Nagaraj, N. Aging and cardiovascular complexity: effect of the length of RR tachograms. *PeerJ* **4**, e2755 (2016).
36. Chouchou, F. & Desseilles, M. Heart rate variability: A tool to explore the sleeping brain? *Front. Neurosci.* **8**, 1–9 (2014).
37. Ziv, J. & Lempel, A. Compression of Individual Sequences via Variable-Rate Coding. *IEEE Trans. Inf. Theory* **24**, 530–536 (1978).
38. Speidel, U., Titchener, M. & Yang, J. How well do practical information measures estimate the Shannon entropy? *Proc. 5th Int. Conf. Information, Commun. Signal Process.* 861–865 (2006).
39. Titchener, M. R. Deterministic computation of complexity, information and entropy. *Inf. Theory, 1998. Proceedings. 1998 IEEE Int. Symp.* **22**, 326 (1998).
40. Cosma Shalizi. Complexity, Entropy and the Physics of gzip. Available at, http://bactra.org/notebooks/cep-gzip.html (Accessed: 12th November 2018) (2003).
41. Chen, C. H. *et al.* Complexity of Heart Rate Variability Can Predict Stroke-In-Evolution in Acute Ischemic Stroke Patients. *Sci. Rep.* **5**, 1–5 (2015).
42. Zhang, Y., Wei, S., Di Maria, C. & Liu, C. Using Lempel–Ziv Complexity to Assess ECG Signal Quality. *J. Med. Biol. Eng.* **36**, 625–634 (2016).
43. Valenza, G. *et al.* Complexity Variability Assessment of Nonlinear Time-Varying Cardiovascular Control. *Sci. Rep.* **7**, 1–15 (2017).
44. Thijs, V. Atrial fibrillation detection fishing for an irregular heartbeat before and after stroke. *Stroke* **48**, 2671–2677 (2017).
45. Matthews, G. D. K., Guzadhur, L., Grace, A. & Huang, C. L.-H. Nonlinearity between action potential alternans and restitution, which both predict ventricular arrhythmic properties in $Scn5a^{+/-}$ and wild-type murine hearts. *J. Appl. Physiol.* **112**, 1847–1863 (2012).
46. Matthews, G. D. K., Guzadhur, L., Sabir, I. N., Grace, A. A. & Huang, C. L. H. Action potential wavelength restitution predicts alternans and arrhythmia in murine $Scn5a^{+/-}$ hearts. *J. Physiol.* **591**, 4167–4188 (2013).
47. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Available at, https://www.r-project.org/ (2016).

Acknowledgements

Author Contributions

V.A. planned the study, undertook analysis and prepared the manuscript. J.A.F. undertook analysis and prepared the manuscript. A.D. provided technical expertise on complexity analysis and revised draft manuscript. M.B. provided clinical input and reviewed all drafts of the manuscript. C.L.H. provided input of project design and reviewed all drafts of the manuscript. C.M.M. performed data collection, provided clinical input and reviewed all drafts of the manuscript and secured funding for the project. K.J. planned the study, reviewed all drafts of the manuscript and secured funding for the project.

Additional Information

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.