



Sampling frequency influences sample entropy of kinematics during walking

Peter C. Raffalt^{1,2} · John McCamley³ · William Denton⁴ · Jennifer M. Yentes⁴ 

Received: 7 March 2018 / Accepted: 21 October 2018

© International Federation for Medical and Biological Engineering 2018

Abstract

Sample entropy (SaEn) has been used to assess the regularity of lower limb joint angles during walking. However, changing sampling frequency and the number of included strides can potentially affect the outcome. The present study investigated the effect of sample frequency and the number of included strides on the calculations of SaEn in joint angle signals recorded during treadmill walking. Eleven subjects walked at their preferred walking speed for 10 min, and SaEn was calculated on sagittal plane hip, knee, and ankle angle signals extracted from 50, 100, 200, 300, and 400 strides at sampling frequencies of 60, 120, 240, and 480 Hz. Increase in sampling frequency decreased the SaEn significantly for the three joints. The number of included strides had no effect on the SaEn calculated on the hip joint angle and only limited effect on the SaEn calculated on the knee and ankle joint signals. The present study suggests that the number of data points within each stride to a greater extent determines the size of the SaEn compared to the number of strides and emphasizes the use of a fixed number of data points within each stride when applying SaEn to lower limb joint angles during walking.

Keywords Gait · Dynamics · Regularity · Joint angles · Methodology

1 Introduction

The concept of information entropy was first introduced by Claude Shannon in 1948 as a method of quantifying the amount of information in communication [16]. The original idea was based on the probability of information coming from various sources and later became the foundation for various versions of entropy measures used to quantify predictability (e.g., approximate entropy, sample entropy (SaEn), fuzzy

entropy). SaEn was introduced by Richmann and Moorman [14] and is the most used version in biomechanical research to quantify predictability. It has been used to investigate standing [9] and sitting [17] balance, gait impairment in various individuals [2, 13, 19, 22], sport performance [12], and force production [4, 18].

Compared to other widely used nonlinear methods applied to gait data (e.g., the largest Lyapunov exponent, correlation dimension and detrended fluctuation analysis), SaEn has gained increasing interest in the last two decades [23]. While the aforementioned methods have been associated with both methodological and interpretational challenges, the computation and interpretation of SaEn is relative straightforward [5, 7, 8, 23]. We have previously shown that when applied appropriately, SaEn constitutes a valuable tool for characterizing the dynamics of human movements in terms of predictability [23, 24]. In gait analysis, SaEn has been used on various types of discrete and continuous data such as stride-to-stride time intervals [2, 21], stride length and step width fluctuations [21], joint angles [1, 13, 19, 22], and trunk accelerations [10, 11, 15]. Although deterministic in nature, stride characteristics exhibit a stochastic behavior in contrast to the pseudo-periodic behavior of joint angles and trunk accelerations. Furthermore, the stride characteristic variables are discrete

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11517-018-1920-2>) contains supplementary material, which is available to authorized users.

✉ Jennifer M. Yentes
jyentes@gmail.com

¹ Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité—Universitätsmedizin Berlin, Berlin, Germany

² Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark

³ MORE Foundation, Phoenix, AZ, USA

⁴ Center for Research in Human Movement Variability, Department of Biomechanics, University of Nebraska Omaha, Omaha, NE, USA

variables and the joint angles and trunk acceleration signals are continuous variables. This difference could have important methodological implications.

During treadmill walking, stride time intervals and joint angles can be extracted from motion capture based kinematic data. Altering the sampling frequency will increase the detail level for both variables. For the stride time interval, changing the sampling frequency from 60 Hz to 240 Hz will change the detection level from 0.0167 to 0.0042 seconds. Equally, this change in sampling frequency will increase the number of data points fourfold within each stride for the joint angle signals. We have previously shown that data length affects the outcome of SaEn analysis when using stride-to-stride time intervals [23, 24]. However, it is unknown how changing the number of data points within each stride will affect the SaEn calculated on cyclic data like joint angles. Furthermore, since the use of SaEn in joint angle data should aim at investigating the regularity within the stride-to-stride fluctuations and not just within a single stride, it is equally important to investigate the effect of changing the number of included strides. Thus, the purpose of the present study was to investigate the effect of sample frequency and the number of included strides on the calculations of sample entropy in joint angle signals recorded during treadmill walking.

2 Methods

2.1 Subjects

Eleven healthy subjects (8 males) with mean (SD) age of 24 years (3), body mass of 74.7 kg (12.4), and body height of 174.2 cm (6.9) were recruited for the present study. All subjects were without lower limb injuries and neurological disorders that could affect the gait pattern. Informed consent was obtained from all individual participants included in the study. The study was approved by the Institutional Review Board of the University of Nebraska Medical Center and the study was carried out in accordance with the approved guidelines.

2.2 Experimental setup

A 12 high-speed camera system operating at 480 Hz (Motion Analysis Corp., Santa Rosa, CA) recorded the three-dimensional position of 27 reflective markers placed bilaterally on: (1) anterior superior iliac spines, (2) posterior superior iliac spines, (3) greater trochanters, (4) midlateral thighs, (5) lower front thighs, (6) lateral knees, (7) tibial tubercles, (8) lower lateral shanks, (9) lateral ankles, (10) top of the second metatarsophalangeal (MTP) joints, (11) posterior heels, (12) lateral fifth MTPs, and (13) lateral calcanei. An additional single marker was placed on the sacrum. The subjects were

walking at their preferred walking speed on a treadmill (AMTI force-sensing tandem treadmill). Prior to the data collection, the preferred walking speed of the subjects were established by increasing and decreasing the speed of the treadmill above and below what the subjects reported as being comfortable. The average speed was 1.17 ± 0.23 m/s. The subject walked for 10 min during which data was collected.

2.3 Data analysis

Sagittal hip, knee, and ankle joint angles were extracted from the kinematic data using Visual 3D (c-Motion, Inc., Germantown, MD). The joint angles from 400, 300, 200, 100, and 50 consecutive strides were extracted. For each data length, the original time series with a sampling frequency of 480 Hz were additionally down sampled to 240, 120, and 60 Hz. SaEn was calculated from the joint angle time series based on the algorithm presented by Richman and Moorman [14] and defined as the negative logarithm for the conditional probability that a series of data points within a certain distance, m , would be repeated within the distance $m+1$ (Eq. 1).

$$SaEn(m, r, N) = -\ln \left[\frac{A^{m+1}(r)}{B^m(r)} \right] \quad (1)$$

where N is the number of data points in the time series, A is the number of similar vector lengths ($m+1$) falling within a relative tolerance limit (r times standard deviation of the time series) and B is the number of similar vector lengths (m) falling within the tolerance limit [23]. To investigate the effect of parameter choice, SaEn calculations were performed with $m=2$ and 3 and $r=0.1, 0.15, 0.2, 0.25$, and 0.3. The results from the analysis using $m=2$ and $r=0.2$ is presented below and the remaining results are presented in a [supplementary material](https://digitalcommons.unomaha.edu/) available at the University of Nebraska, Omaha Digital Commons (<https://digitalcommons.unomaha.edu/>).

2.4 Statistics

The applied statistical models assumed independency of observation, normal distribution of the residuals, and homogeneity of variances (homoscedasticity). The normal distribution was confirmed by the Shapiro-Wilk test and the homogeneity was confirmed by the Brown-Forsythe test. To investigate if there was an overall effect of sampling frequency and number of included strides on the SaEn calculated on each of the three joint angles, a two-way ANOVA for repeated measures was applied to each joint angle with sampling frequency and number of included strides as independent variables. In case of an overall significant effect of the independent variables or the

interaction (sampling frequency \times strides), a Holm Sidak post hoc test was applied to investigate differences in sampling frequency and number of strides. To investigate if there was an overall effect of joint on the SaEn, a one-way ANOVA for repeated measures was applied to the SaEn calculated from 400 strides with a sampling rate of 480 Hz. This combination of sampling frequency and number of strides was chosen to include as much information as possible in the comparison. In case of an overall significant effect of joint, a Holm Sidak post hoc test was applied to investigate between joint differences. Level of significance was set at 0.05. All statistical analyses were performed in Sigmaplot (Systat Software, Inc. 2014, version 13.0, Germany).

3 Results

There was an overall significant effect of sampling frequency on the SaEn for the hip joint angle ($p < 0.001$) (Fig. 1). The post hoc tests revealed that with each increase in sampling frequency, the SaEn decreased significantly ($p < 0.001$ for all comparisons). There was an overall significant effect sampling frequency ($p < 0.001$) and of the number of included strides ($p = 0.049$) on the SaEn for the knee joint angle (Fig. 1). The SaEn decreased significantly with each increase in sampling frequency. In contrast, the post hoc test revealed no differences in SaEn between the different numbers of strides. There was an overall significant effect of both sampling frequency ($p < 0.001$) and of the number of strides included ($p < 0.001$) on the SaEn for the ankle joint angle (Fig. 1). The post hoc tests revealed that with each increase in sampling frequency the SaEn decreased significantly ($p < 0.001$ for all comparisons). Furthermore, SaEn calculated from 50 and 100 strides was significantly lower compared to SaEn calculated from 300 and 400 strides. Additionally, SaEn calculated from 200 strides was significantly lower compared to SaEn calculated from 400 strides.

When comparing SaEn from the three joints calculated from 400 strides at 480 Hz (fig. 1), there was an overall significant effect of joint ($p < 0.001$). The post hoc test revealed the SaEn from the ankle joint was significantly higher compared to the SaEn of the two more proximal joints ($p < 0.001$ in both cases). There was no significant difference between the SaEn from the knee and hip joint angles.

The results presented in the [supplementary material](#) indicated the same overall tendency as presented above. For all input parameter combinations, there was a significant effect of sampling frequency with SaEn decreasing considerable with increase in sampling frequency. However, the observed overall effect of the number of included strides resulted in relative small absolute changes in the SaEn.

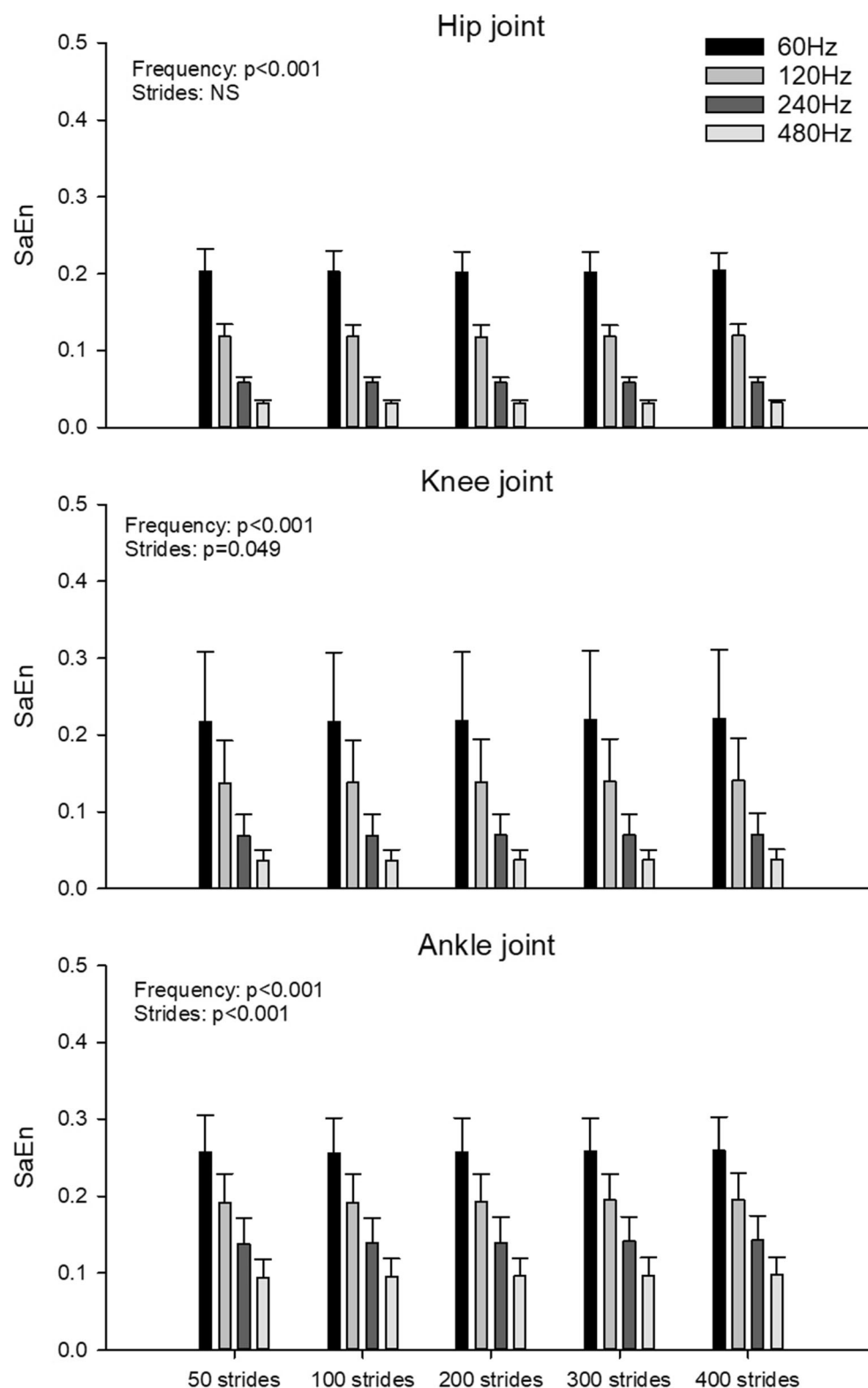
4 Discussion

The present study aimed at investigating the effect of sample frequency and the number of included strides when quantifying regularity using SaEn on lower limb joint angles during walking. The main result was that while SaEn decreased considerably when sampling frequency was increased, changing the number of included strides seemed to have limited effect of the SaEn. These effects did not differ substantially between the three joints investigated in the present study.

We have previously observed considerable effects of changes in input parameters when using SaEn to quantify regularity in step-to-step time characteristics during walking [23, 24]. Especially, time series including less than 200 steps exhibited low parameter consistency [23]. In contrast, the number of included strides had limited effect on the SaEn in the present study. Assuming that the average stride time was 1 second, the lowest number of data points in the present study was ($60 \text{ Hz} \times 50 \text{ strides} \times 1 \text{ second}$) 3000. Increasing the number of strides to 400 but keeping the same sampling frequency resulted in ($60 \text{ Hz} \times 400 \text{ strides} \times 1 \text{ second}$) 24,000 data points, just as keeping the number of strides at 50 and increasing the sampling frequency to 480 Hz did ($480 \text{ Hz} \times 50 \text{ strides} \times 1 \text{ second}$). While the first example had limited effect of the SaEn, the other decreased the SaEn substantially. These two examples of increasing the total number of data points resulted in two different changes in SaEn. This illustrates that the number of data points within each stride affects SaEn to a greater extent than the number of strides. It could be speculated that the decrease in SaEn with increase in sampling frequency could be related to the nature of the algorithm. Since an increase in sampling frequency does not change the standard deviation of the investigated time series, the applied critical limit will not change. However, as the sampling frequency increases the spatial distance between adjacent data points will decrease which will increase the number of vectors within each stride cycle. This will increase the likelihood of identifying vectors of $m+1$ that falls within vectors of m in subsequent strides. In contrast, increasing the number of included stride cycles without increasing the sampling frequency will not increase the number of vectors within each stride cycle and the likelihood of identifying similar vectors will not change significantly.

The present study suggests that data from various populations can be compared even if the number strides is not the same as long as the number of data points within each stride is kept the same. This supports the use of SaEn on joint angles during walking in patients groups or frail individuals incapable to complete long recording sessions which is needed when assessing regularity in stride-to-stride time intervals [23, 24]. Some of the previous studies applying SaEn to lower limb joint angles recorded during walking have used a range of sampling frequencies from 60 Hz [22] to 200 Hz [19]. In light of the results from the present study, SaEn values from studies

Fig. 1 SaEn calculated on the hip joint angle (top graph), knee joint angle (middle graph), and ankle joint angle (bottom graph) for time series including 50, 100, 200, 300, and 400 strides and sampled at 60, 120, 240, and 480 Hz



with different sampling frequencies should not be compared directly. However, the present study cannot recommend the use of a specific sampling frequency. Therefore, the sampling frequency of future studies should be determined a priori

based on the research design and appropriate power spectral analysis.

Our previous studies have revealed some input parameter inconsistency when applying SaEn to stride-to-stride time

interval time series with the same range of input parameters as in the present study [23, 24]. The present study observed an acceptable input parameter consistency across the different sampling frequencies and numbers of included strides. Thus, regardless of the choice of input parameters, the SaEn changed in a similar manner across sampling frequencies and number of strides. Additionally, the between-joint differences in SaEn did not differ across different input parameters.

While the present study elucidates important methodological aspects of applying SaEn to lower limb joint angles recorded during walking, it also indicates potential differences in these methodological aspects when comparing SaEn applied to continuous time series (e.g., joint angles, center of mass trajectories) compared to time series including discrete events (e.g., stride-to-stride time interval time series). Thus, future studies should investigate this topic further.

SaEn is one of several nonlinear methods, which has received increasing interest in both basic and applied gait research in recent years. While methods like detrended fluctuation analysis, the largest Lyapunov exponent and correlation dimension offers quantification various characteristics of the dynamics of human locomotion, the implementation and interpretation possess inherent challenges which has led to inconsistent observations and conclusions in the literature [3, 5, 6, 8, 20]. The interpretation and calculation of SaEn is relative straightforward once a few methodological aspect related to the input parameter consistency are accounted for [23, 24]. In light of the present study, we do not believe that the methodological constraints mentioned should prevent the application of SaEn to continuous cyclic signals recorded during human locomotion.

In conclusion, the present study observed that while increasing the sampling frequency significantly reduced the SaEn of joint angle trajectories during walking increasing the number of included stride cycles had limited effect. Based on these observations, we recommend that when SaEn is applied to lower limb joint angles recorded during walking, the number of data points within each stride is fixed across all subjects even if the number of recorded strides differs. This could be done by keeping the sampling frequency fixed and the walking speed fixed across subjects. Alternatively, the recorded time series are resampled to a fixed number of data points within each stride.

Acknowledgements The authors would like to thank Sidney Baudendistel and Kristen Watson for their assistance during data collection.

Author's contributions PCR, JM, and JMY conceived and designed the study, JM, WD and JMY performed the experiment; PCR and WD analyzed the data; and PCR, JM, WD, and JMY interpreted the results, drafted, edited, and approved the manuscript.

Funding Funding for this project was provided by the National Institutes of Health (P20 GM109090).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflicts of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

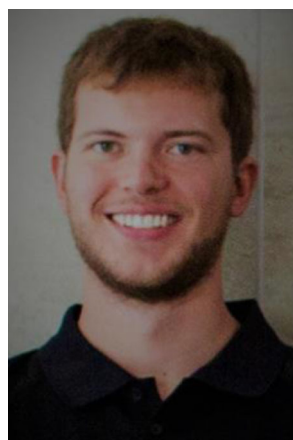
References

1. Alkjaer T, Raffalt P, Petersen NC, Simonsen EB (2012) Movement behavior of high-heeled walking: how does the nervous system control the ankle joint during an unstable walking condition? *PLoS One* 7:e37390. <https://doi.org/10.1371/journal.pone.0037390>
2. Alkjaer T, Raffalt PC, Dalsgaard H, Simonsen EB, Petersen NC, Bliddal H, Henriksen M (2015) Gait variability and motor control in people with knee osteoarthritis. *Gait Posture* 42:479–484. <https://doi.org/10.1016/j.gaitpost.2015.07.063>
3. Bruijn SM, Meijer OG, Rispens SM, Daffertshofer A, van Dieën JH (2012) Letter to the editor: “sensitivity of the Wolf’s and Rosenstein’s algorithms to evaluate local dynamic stability from small gait data sets”. *Ann Biomed Eng* 40:2505–2506
4. Chow JW, Stokic DS (2016) Variability, frequency composition, and temporal regularity of submaximal isometric elbow flexion force in subacute stroke. *Exp Brain Res* 234:3145–3155. <https://doi.org/10.1007/s00221-016-4712-7>
5. Cignetti F, Decker LM, Stergiou N (2012) Sensitivity of the Wolf’s and Rosenstein’s algorithms to evaluate local dynamic stability from small gait data sets. *Ann Biomed Eng* 40:1122–1130. <https://doi.org/10.1007/s10439-011-0474-3>
6. Cignetti F, Decker LM, Stergiou N (2012) Sensitivity of the Wolf’s and Rosenstein’s algorithms to evaluate local dynamic stability from small gait data sets: response to commentaries by Bruijn et al. *Ann Biomed Eng* 40:2507–2509
7. Delignieres D, Torre K, Bernard PL (2011) Transition from persistent to anti-persistent correlations in postural sway indicates velocity-based control. *PLoS Comput Biol* 7:e1001089. <https://doi.org/10.1371/journal.pcbi.1001089>
8. Dingwell JB, Cusumano JP (2010) Re-interpreting detrended fluctuation analyses of stride-to-stride variability in human walking. *Gait Posture* 32:348–353. <https://doi.org/10.1016/j.gaitpost.2010.06.004>
9. Donker SF, Roerdink M, Greven AJ, Beek PJ (2007) Regularity of center-of-pressure trajectories depends on the amount of attention invested in postural control. *Exp Brain Res* 181:1–11. <https://doi.org/10.1007/s00221-007-0905-4>
10. Kosse NM, Vuillerme N, Hortobagyi T, Lamothe CJ (2016) Multiple gait parameters derived from iPod accelerometry predict age-related gait changes. *Gait Posture* 46:112–117. <https://doi.org/10.1016/j.gaitpost.2016.02.022>
11. Lamothe CJ, Ainsworth E, Polonski W, Houdijk H (2010) Variability and stability analysis of walking of transfemoral amputees. *Med Eng Phys* 32:1009–1014. <https://doi.org/10.1016/j.medengphys.2010.07.001>

12. Preatoni E, Ferrario M, Dona G, Hamill J, Rodano R (2010) Motor variability in sports: a non-linear analysis of race walking. *J Sports Sci* 28:1327–1336. <https://doi.org/10.1080/02640414.2010.507250>
13. Rhea CK, Wutzke CJ, Lewek MD (2012) Gait dynamics following variable and constant speed gait training in individuals with chronic stroke. *Gait Posture* 36:332–334. <https://doi.org/10.1016/j.gaitpost.2012.03.014>
14. Richman JS, Moorman JR (2000) Physiological time-series analysis using approximate entropy and sample entropy. *Am J Phys Heart Circ Phys* 278:H2039–H2049
15. Rispens SM, Van Dieën JH, Van Schooten KS, Cofre Lizama LE, Daffertshofer A, Beek PJ, Pijnappels M (2016) Fall-related gait characteristics on the treadmill and in daily life. *J Neuroeng Rehabil* 13:12. <https://doi.org/10.1186/s12984-016-0118-9>
16. Shannon C (1948) A mathematical theory of communication. *Bell Syst Tech J* 27:379–423
17. Sondergaard KH, Olesen CG, Sondergaard EK, de Zee M, Madeleine P (2010) The variability and complexity of sitting postural control are associated with discomfort. *J Biomech* 43:1997–2001. <https://doi.org/10.1016/j.jbiomech.2010.03.009>
18. Svendsen JH, Madeleine P (2010) Amount and structure of force variability during short, ramp and sustained contractions in males and females. *Hum Mov Sci* 29:35–47. <https://doi.org/10.1016/j.humov.2009.09.001>
19. Terada M, Bowker S, Thomas AC, Pietrosimone B, Hiller CE, Rice MS, Gribble PA (2015) Alterations in stride-to-stride variability during walking in individuals with chronic ankle instability. *Hum Mov Sci* 40:154–162. <https://doi.org/10.1016/j.humov.2014.12.004>
20. Terrier P, Reynard F (2018) Maximum Lyapunov exponent revisited: long-term attractor divergence of gait dynamics is highly sensitive to the noise structure of stride intervals. *Gait Posture*. <https://doi.org/10.1016/j.gaitpost.2018.08.010>
21. Thomas KS, Russell DM, Van Lunen BL, Colberg SR, Morrison S (2017) The impact of speed and time on gait dynamics. *Hum Mov Sci* 54:320–330. <https://doi.org/10.1016/j.humov.2017.06.003>
22. Wurdeman SR, Myers SA, Stergiou N (2014) Amputation effects on the underlying complexity within transtibial amputee ankle motion. *Chaos* 24:013140. <https://doi.org/10.1063/1.4869580>
23. Yentes JM, Hunt N, Schmid KK, Kaipust JP, McGrath D, Stergiou N (2013) The appropriate use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng* 41:349–365. <https://doi.org/10.1007/s10439-012-0668-3>
24. Yentes JM, Denton W, McCamley J, Raffalt PC, Schmid KK (2018) Effect of parameter selection on entropy calculation for long walking trials. *Gait Posture* 60:128–134. <https://doi.org/10.1016/j.gaitpost.2017.11.023>



John McCamley received his PhD in Bioengineering from the University of Bologna in 2013. He is currently director of the Human Movement Laboratory, MORE Foundation Phoenix, Arizona.



William Denton is a laboratory technician in the Center for Human Movement Variability (MOVCENTR), Department of Biomechanics at the University of Nebraska at Omaha. He has degrees in Biological Systems Engineering and Biomechanics.



Jennifer M. Yentes is an Assistant Professor in the Center for Human Movement Variability (MOVCENTR), Department of Biomechanics at the University of Nebraska at Omaha. Her research covers human movement variability.



Peter C. Raffalt is a Postdoctoral Researcher at the Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité—Universitätsmedizin Berlin in Germany. His research covers human movement variability.