The effect of run duration, gait variable and Lyapunov exponent algorithm on the inter-session reliability of local dynamic stability in healthy young people

KENNERLEY, Adam S, DUNN, Marcus <http://orcid.org/0000-0003-3368-8131>, MIDDLETON, Kane, WEBSTER, Kate E. and WHEAT, Jonathan <http://orcid.org/0000-0002-1107-6452>

Human Movement Science 99 (2025) 103325

Contents lists available at ScienceDirect

Human Movement Science

journal homepage: www.elsevier.com/locate/humov

Full Length Article

The effect of run duration, gait variable and Lyapunov exponent algorithm on the inter-session reliability of local dynamic stability in healthy young people

Adam S. Kennerley [a,c,*], Marcus Dunn [b], Kane Middleton [c], Kate E. Webster [c], Jonathan Wheat [d]

[a] *Sports Engineering Research Group, Sport and Physical Activity Research Centre, Sheffield Hallam University, Sheffield, UK*
[b] *School of Energy, Geoscience, Infrastructure and Society, Institute for Life and Earth Sciences, Heriot-Watt University, Edinburgh, UK*
[c] *Sport, Performance, and Nutrition Research Group, School of Allied Health, Human Services and Sport, La Trobe University, Melbourne, Victoria, Australia*
[d] *Department of Sport Sciences, Nottingham Trent University, Nottingham, UK*

A R T I C L E I N F O

Keywords:
Nonlinear analysis
Rehabilitation
Gait analysis
Running
Injury

A B S T R A C T

Local dynamic stability (LDS) of gait has been used to differentiate between healthy and injured populations, establishing its potential as an indicator of healthy gait and a new objective measure to assess gait function following injury. For LDS to be a reliable assessment tool of healthy gait progression during rehabilitation, it must provide consistent and sensitive inter-session measures. Methodological factors such as trial duration, gait variable, and Lyapunov Exponent (LyE) algorithm can influence LDS estimation and its reliability. Young people are a high-risk population for sport-related injuries, and running is a key activity during rehabilitation and is regularly assessed. Therefore, the effects of run duration, gait variable, and LyE algorithm choice on the reliability and sensitivity of inter-session LDS measures in young people were investigated. Sixteen healthy participants ran on a treadmill on two separate sessions (difference of 7 ± 5 days). LDS was calculated using both the Rosenstein and Wolf algorithm for durations of 1-, 2-, 3-, 4- and 5-min of knee flexion angle and medio-lateral acceleration of the pelvis and thorax from each session. The relative and absolute reliability between sessions was calculated using the intraclass correlation coefficient and standard error of measurement. The sensitivity of inter-session LDS change was quantified by the minimal detectable change. Results showed that longer run durations produced higher relative reliability and a minimum run duration of 4 min is recommended to achieve moderate-to-good inter-session reliability across all gait variables and LyE algorithms. However, shorter durations of 2–3 min may still be sufficient when using medio-lateral pelvis acceleration or knee flexion angle, particularly with the Rosenstein algorithm, which also improves sensitivity to change. These findings provide practical guidance for methodological choices when calculating LDS in young people during running and support their

Abbreviations: λ, Lyapunov exponent (Unit of information = natural logarithms per second); 95 %CI, 95 % confidence intervals; ICC, Intraclass correlation coefficient; KFA, Knee flexion angle; KFAdom, Knee flexion angle of the dominant leg; KFAndom, Knee flexion angle of the non-dominant leg; LDS, Local dynamic stability; LyE, Lyapunov exponent; MDC, Minimal detectable change; PelvisML, Pelvis segment centre of mass medio-lateral linear accelerations; PRS, Preferred running speed; PTS, Preferred transition speed; SEM, Standard error of measurement; ThoraxML, Thorax segment centre of mass medio-lateral linear accelerations.

* Corresponding author at: Sports Engineering Research Group, Sport and Physical Activity Research Centre, Sheffield Hallam University, Sheffield, UK.

https://doi.org/10.1016/j.humov.2025.103325
Received 4 September 2024; Received in revised form 6 December 2024; Accepted 24 January 2025
Available online 30 January 2025
0167-9457/

potential use as reliable tools for monitoring gait function and tracking rehabilitation progress in young people following injury.

1. Introduction

Local dynamic stability (LDS) of gait is a construct with the potential to objectively assess movement function and reflects a system's ability to manage small internal perturbations derived from sensory inputs or neuromuscular noise (Bruijn et al., 2013; Full, 2002). Thus, a lack of LDS may lead to undesired movement strategies and potential injury, and may also be indicative of lower limb injury (Strongman & Morrison, 2020). Differences in LDS of walking gait between healthy and injured individuals have been reported (Afonso De Oliveira et al., 2019; Mahmoudian et al., 2016; Moraiti et al., 2010), highlighting LDS as a potential indicator of healthy gait function. Furthermore, changes in LDS of walking have been seen over time during rehabilitation from injury (Afonso De Oliveira et al., 2019; Castiglia et al., 2024; Hilfiker et al., 2013), emphasising that LDS measures might provide additional insight into an athlete's recovery and readiness to return to sport. However, despite return to running being a key goal when returning to sport and its monitoring a common clinical assessment (Bramah et al., 2024; Buckthorpe et al., 2020; Myer et al., 2006), there is a paucity of studies investigating LDS during running. Furthermore, young people are a high-risk population for sport-related injury (Conn et al., 2003; Kisser & Bauer, 2012). The impact of injury affects both their acute (physical activity participation) and chronic (early osteoarthritis) physical health (Ardern et al., 2014; Whittaker et al., 2015), and so a successful return to physical activity is paramount. Therefore, strategies to provide measures of healthy movement in a sporting context i.e. running, for young people would be impactful. Furthermore, there is a lack of studies investigating LDS in young people specifically.

LDS of gait can be estimated by calculating the largest Lyapunov exponent (LyE or λ) of a time series (e.g., joint angle or segment acceleration). The LyE of a time series is an assessment of the divergence in movement trajectories from one stride to the next (Stergiou, 2016). Too much or too little divergence suggests a system that lacks adaptability to deal with perturbations, producing movement patterns for gait kinematics that have the potential to cause falls or injury (Bartlett et al., 2007; Stergiou & Decker, 2011). Commonly assessed running-gait kinematic variables in the estimation of LDS are knee flexion angle (KFA) and trunk movement (Hunter et al., 2023; Winter et al., 2023). Further, medio-lateral trunk movement may be a more specific choice as it is crucial for gait stability, relying on neuromuscular control, while other planes are stabilized by the musculoskeletal system (Allet et al., 2012). Additionally, KFA adjustment and neuromuscular control are also vital for maintaining joint stability particularly after knee injury (Schrijvers et al., 2021; Sharifi & Shirazi-Adl, 2021).

The investigation into the reliability of LDS measures during running would inform the potential application of LDS as a repeated measure of running function e.g. in young athletes who are rehabilitating to sport. It has been suggested that methodological choices such as LyE algorithm, trial duration, and gait variable affect the reliability of LyE calculation (Kang & Dingwell, 2006; Raffalt, 2018; Reynard & Terrier, 2014). For example, Raffalt (2018) assessed the inter-session reliability of LDS during walking using knee flexion angle (KFA) and sacrum position, using two LyE algorithms: Rosenstein (Rosenstein et al., 1992) and Wolf (Wolf et al., 1985). LDS estimated using the Rosenstein algorithm demonstrated good-to-excellent reliability while the Wolf algorithm estimates of LDS produced poor reliability. The LDS of KFA was more reliable than the LDS of sacrum position, regardless of algorithm. Conversely, other studies have shown that good-to-excellent inter-session reliability of LDS can be achieved with the Wolf algorithm (Nazary-Moghadam et al., 2020) and also with trunk variables (i.e., acceleration of the sacrum and sternum) (Ekizos et al., 2018; Fohrmann et al., 2022). Conflicting results such as these make methodological choices difficult when the goal is to reliably measure LDS. However, there is agreement that longer trials or analysing more strides improves the reliability of LDS measures (Kang & Dingwell, 2006; Reynard & Terrier, 2014). However, identifying a minimum trial duration, which elicits good reliability, would help minimise any potential burden placed on participants e.g. injured athletes undergoing repeat LDS assessments during rehabilitation.

Determining the effect of trial duration, gait variable, and LyE algorithm on the between-session sensitivity and reliability of LDS during running in young people, would not only add to the growing body of knowledge that focuses on LyE calculation parameter choices; but would inform future calculations of repeat LDS measures in young populations during running. Therefore, the aim of this study is to assess how run duration, gait kinematic variable, and LyE algorithm selection influence the inter-session reliability and sensitivity of LDS in a young population during running.

2. Methods

2.1. Participants

Sixteen healthy young people who competed in sports at least once a week, participated in the study after giving informed consent or informed assent with parental consent where required; all procedures were approved by both Sheffield Hallam University and La Trobe University's Research Ethics Committee (ER34023326). After examining LyE distribution boxplots one participant was removed as an outlier as they consistently produced values of >1.5 the inter quartile range and was also identified as the only athlete who competed in a sport that did not involve running (i.e., dancing). The remaining 15 participants (8 female, 7 male; age = 15 ± 2 years old; height = 175 ± 8 cm; mass = 65.3 ± 10.7 kg) were free from any neuromuscular pathology or injury that could affect their gait. A priori power analysis using G*Power 3 (Faul et al., 2007) was performed based on a previous effect size ICC = 0.6 (Reynard & Terrier, 2014), with a power of 80 % and a significance level of 0.05, and a sample size of 15 participants was considered suitable.

2.2. Protocol

Participants ran on a treadmill (Precor TRM-731, Amer Sports Ltd) for 5.5 min at a self-selected preferred running speed (PRS) and repeated this at a second session at least 48 h later (mean 7 ± 5 days). PRS was determined using a "method of limits" (Treutwein, 1995) design, similar to a previously reported protocol (Dingwell & Marin, 2006). Starting from a preferred walk-run transition speed (PTS) (mean 7.0 ± 0.4 km/h) the treadmill speed was increased by 0.1 km/h every 2 s until the participant reported it was faster than a preferred running speed. Speed was then increased by 0.5 km/h before being decreased by 0.1 km/h every 2 s until the participant reported it was slower than a preferred running speed. The mean of the two self-reported speeds was determined as the PRS (mean 9.0 ± 1.1 km/h). The protocol started at the PTS to ensure a PRS was selected substantially higher than the PTS as running near to the PTS is known to affect LDS (Jordan et al., 2006).

2.3. Procedures

The three-dimensional position of 19 reflective markers, 2 shank (3 marker) and 2 thigh (4 marker) cluster sets attached to the participant (Fig. 1) was recorded using an 8-camera motion capture system (Miqus M3, Qualisys, Göteborg, Sweden) at a sampling frequency of 200 Hz. A 10-s static standing trial was first recorded to model body segments, after which the femur medial epicondyle and tibia apex of medial malleolus markers were removed for the running trials (Fig. 1).

2.4. Data processing

Knee Flexion Angle of the dominant (KFAdom) and non-dominant (KFAndom) limb, and medio-lateral linear accelerations of the pelvis (PelvisML) and thorax (ThoraxML) were determined as gait variables. These were selected as they have been commonly used to assess the LDS of running gait (Hunter et al., 2023; Winter et al., 2023). Furthermore, medio-lateral trunk movement is crucial for gait stability, relying on neuromuscular control, while other planes are stabilized by the musculoskeletal system (Allet et al., 2012). Additionally, KFA adjustment and neuromuscular control are vital for maintaining joint stability after knee injury (Schrijvers et al., 2021; Sharifi & Shirazi-Adl, 2021). All gait variables were calculated using raw marker positional data, as filtering can affect the calculation of LyE by removing vital information about the system (Raffalt et al., 2020). Body segments were modelled using a six

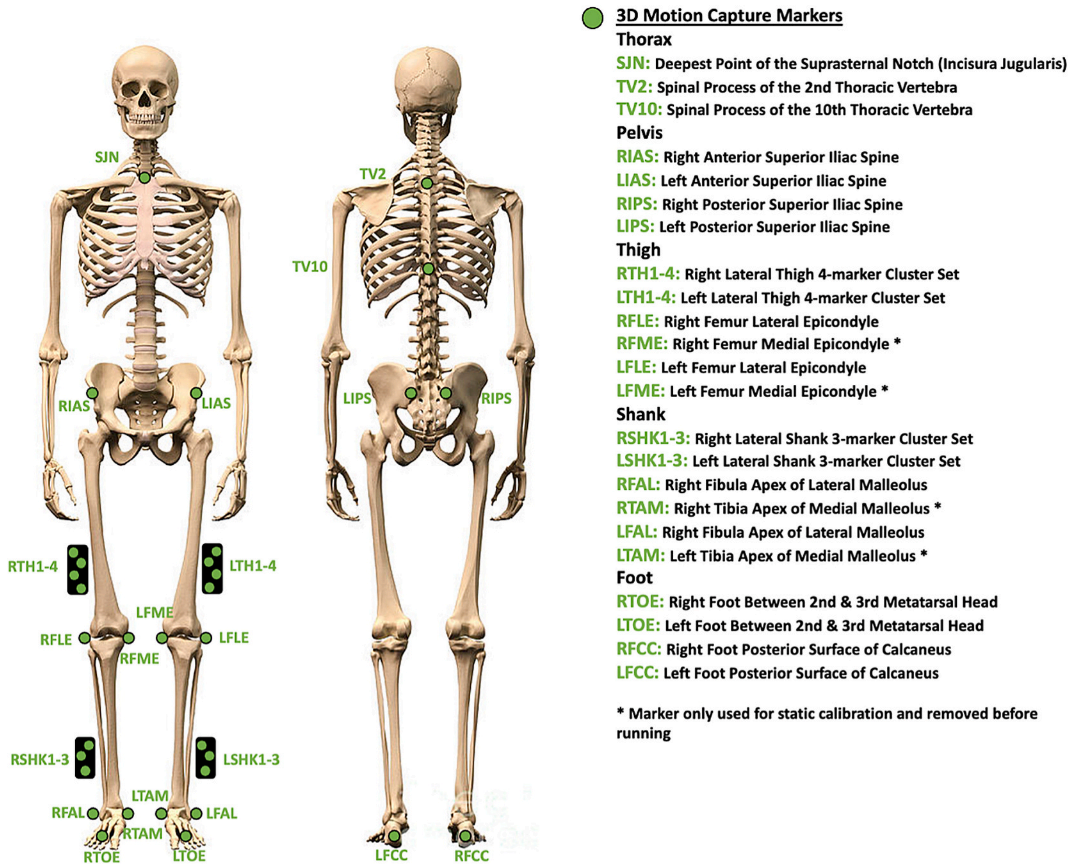

Fig. 1. Locations of the 19 reflective markers, shank and thigh cluster sets used for 3D motion capture analysis.

degrees of freedom segment optimisation algorithm provided by Visual3D (C-Motion Inc., USA). A 3-marker model (Armand et al., 2014) was used for the thorax and the Visual3D 'CODA' model for the pelvis. PelvisML and ThoraxML were defined as the medio-lateral linear accelerations of the respective segment centre of mass and were exported from Visual3D. KFA was defined as a Cardan angle from the transformation of the shank segment to the thigh segment. The dominant limb was selected as the leg with which participants preferred to strike a football (van Melick et al., 2017). Time series data were down sampled to 50 Hz for computational efficiency during the subsequent data analyses and was deemed adequate by examining spectral density plots. The first 30 s of each time series was cropped to remove any gait-initiation effects, after which the remaining 5-min trial was cropped further into 4-, 3-, 2- and 1-min durations each starting from the same first data point (Fig. 2).

For each duration the total number of complete strides were calculated using the heel strikes indicated by the first local minima following each peak KFA in the KFAdom data. The minimum number of strides needed to account for any effect of noise is 50 (Mehdizadeh & Sanjari, 2017) and the 1-min trial duration met this minimum (Table 1). All trial durations for all gait variables were then cropped from the first to the last available heel strike. All gait variables across all participants were then further cropped and resampled to a consistent number of strides and data points corresponding to the minimum number of strides and maximum number of data points calculated for each trial duration (Table 1). Resampling was applied using MATLAB linear interpolation function 'interp1'. This temporal normalisation removes any effect of differing gait speeds, allowing a flexible number of data points per stride while maintaining the temporal structure of each individual stride (Hussain et al., 2021).

2.5. LyE calculation

Each time series was reconstructed in a state space by using the mean time delay ($\tau = 5$) and embedding dimension ($m = 6$) these values remained consistent between sessions and were calculated across all participants, gait variables, and durations. The time delay was calculated from the first minimum of the average mutual information function (Fraser & Swinney, 1986) and the embedding dimension through a global false nearest neighbours analysis (Kennel & Isabelle, 1992).

LyE was calculated from the reconstructed state space using both the Rosenstein and Wolf algorithms. Briefly, both algorithms estimate LyE by calculating the Euclidean distances between nearest neighbouring points on separate trajectories within the state space and calculating the trajectory divergence rate (i.e., rate of separation of these points over time). However, the algorithms differ in their approach. For the Rosenstein algorithm, an initial two neighbouring points were tracked along their separate trajectories within the state space and the natural logarithm of the distance between them was calculated at each point until the end of the time series. This was repeated for all initial neighbouring points in the state space after which the natural log of the average distance located at each similar point was plotted against time. The LyE was then estimated as the slope of a linear fit line on the average log divergence curve. The linear fit line was applied over 22 data points (the mean number of data points in 0.5 strides across all participants and sessions). For the Wolf algorithm, a reference trajectory was selected as the trajectory on which one of an initial two neighbouring points were located. The divergence between trajectories was then calculated after a pre-determined evolution time (22 data points i.e. 0.5 strides), as the natural logarithm of the evolved distance between neighbouring points divided by their original distance. This continued until the distance and angular separation between points exceeded a pre-determined value at which point a new nearest neighbouring point to the reference trajectory point was selected. This process was repeated until the end of the time series. The maximum scaled distance used was equal to the range of the data divided by 10, and the maximum angular separation was 0.2 rad (Wolf et al., 1985). The LyE was then calculated as the mean of all the calculated divergences.

All data processing and LyE calculation was performed in MATLAB (version: 9.13.0 (R2022b), Natick, Massachusetts: Mathworks Inc) using custom scripts adapted from the UNO Biomechanics Nonlinear Analysis Toolbox (UNO Biomechanics, 2023).

2.6. Statistical analysis

The relative and absolute inter-session reliability of LyE, for each run duration, gait variable, and LyE algorithm combination, was

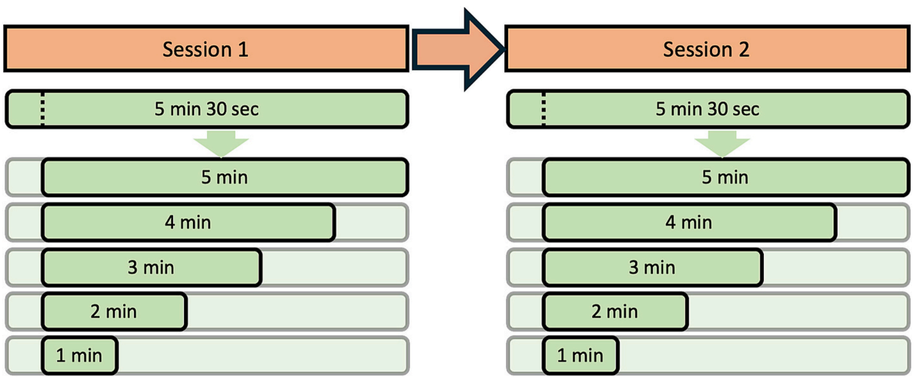

Fig. 2. The duration of gait data collected in each session and how the data was later cropped into 1-, 2-, 3-, 4-, and 5-min durations.

Table 1
The minimum number of strides and maximum number of data points collected across all participants and sessions for each duration.

Duration	Strides	Data Points
1-min	68	2929
2-min	138	5929
3-min	209	8938
4-min	280	11,949
5-min	350	14,954

assessed using Intraclass Correlation Coefficients (ICC) and Standard Error of Measurement (SEM), respectively. $ICC_{(3,1)}$ estimates and their 95 % confident intervals (95 %CI) were based on a single trial measure, absolute-agreement, 2-way mixed-effects model. ICC values were classified as follows: below 0.5 (poor), 0.5 to less than 0.7 (moderate), 0.7 to less than 0.9 (good), and 0.9 or higher (excellent)(Koo & Li, 2016). The square root of the mean square error between sessions across all participants i.e. the variance within individuals, was used to calculate the SEM (Weir, 2005). The minimal detectable change (MDC), which is defined as the minimum change between sessions within an individual that can be considered a real change in LDS i.e. repeat measure sensitivity, was calculated as $1.96 \times \sqrt{2} \times SEM$ (Atkinson & Nevill, 1998). The MDC was then normalised by the mean test-retest LyE value and expressed as a percentage which allows the MDC to evaluate the degree of absolute change within an individual and enables easier comparison across studies. All statistical analysis was performed using SPSS (SPSS Inc., Chicago, IL) with alpha set to 0.05.

3. Results

3.1. Run duration

Durations of 1-, 2-, and 3-min exhibited poor to good relative reliability, with ICCs across both algorithms and all gait variables ranging from 0.14 to 0.79, 0.5 to 0.82, and 0.38 to 0.87, respectively (Fig. 3). Durations of 4- and 5-min exhibited moderate to good relative reliability, with ICCs ranging from 0.62 to 0.85 and 0.61 to 0.84 respectively (Fig. 3). For absolute reliability, durations of 1-, 2-, 3-, 4-, and 5-min showed a maximum SEM of 0.63, 0.51, 0.56, 0.36, and 0.36 λ (Fig. 4); and a maximum MDC of 65 %, 64 %, 76 %, 50 %, and 52 % correspondingly (Fig. 5), across both algorithms and all gait variables.

3.2. Gait variable

Across all durations and algorithms, KFAndom, PelvisML, and ThoraxML exhibited poor to good relative reliability with ICC ranges of 0.49 to 0.87, 0.14 to 0.85, and 0.18 to 0.76 respectively. KFAdom showed moderate to good relative reliability, with ICCs ranging from 0.59 to 0.78 (Fig. 3). From 2-min onwards KFAndom showed moderate to good relative reliability (0.6–0.87) and PelvisML showed good relative reliability (0.72–0.85). From 4-min onwards all gait variables showed moderate to good relative reliability (Fig. 3).

For absolute reliability and measure sensitivity, KFAdom, KFAndom, PelvisML, and ThoraxML exhibited maximum SEMs of 0.16,

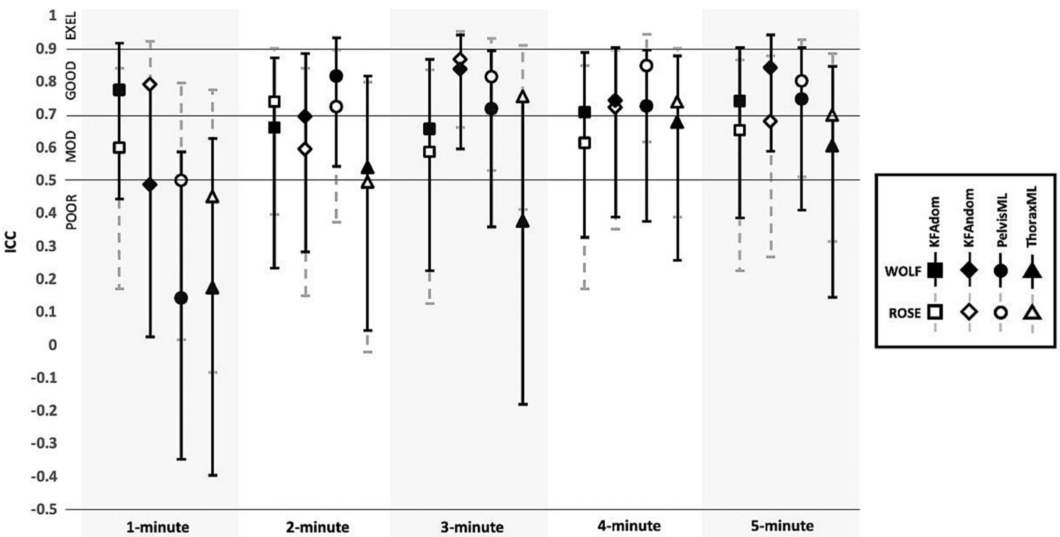

Fig. 3. ICCs and their 95 % confidence intervals for every combination of run duration, gait variable and LyE algorithm.

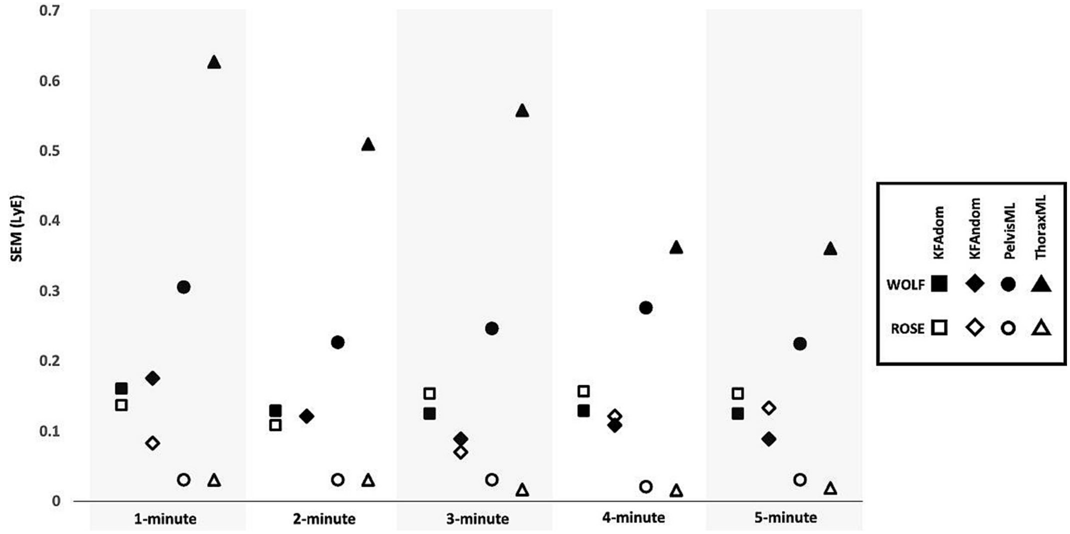

Fig. 4. SEM for every combination of run duration, gait variable and LyE algorithm. Lower values represent greater absolute reliability.

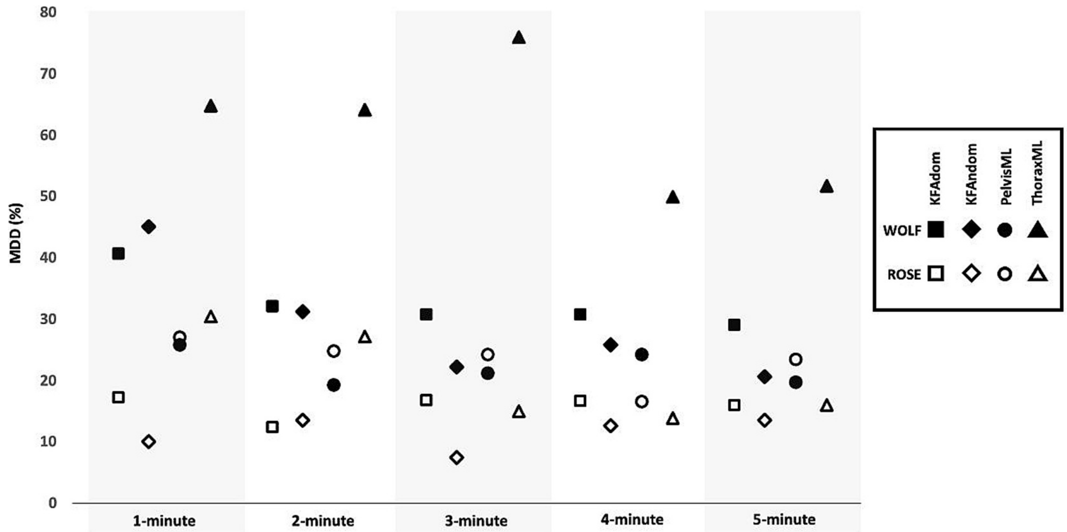

Fig. 5. MDD for every combination of run duration, gait variable and LyE algorithm. Lower values represent greater measure sensitivity to change in LyE.

Table 2
Mean LyE ± standard deviation for each gait variable and run duration for both algorithms.

	1 min	2 min	3 min	4 min	5 min
Rosenstein					
KFAdom	2.21 ± 0.22	2.44 ± 0.21	2.53 ± 0.24	2.61 ± 0.25	2.67 ± 0.26
KFAndom	2.28 ± 0.19	2.48 ± 0.19	2.59 ± 0.20	2.66 ± 0.23	2.73 ± 0.23
PelvisML	0.32 ± 0.06	0.35 ± 0.05	0.36 ± 0.05	0.37 ± 0.06	0.37 ± 0.06
ThoraxML	0.29 ± 0.05	0.32 ± 0.04	0.33 ± 0.03	0.33 ± 0.03	0.34 ± 0.04
Wolf					
KFAdom	1.10 ± 0.32	1.12 ± 0.21	1.14 ± 0.21	1.17 ± 0.23	1.20 ± 0.24
KFAndom	1.08 ± 0.26	1.08 ± 0.24	1.11 ± 0.22	1.17 ± 0.21	1.20 ± 0.21
PelvisML	3.28 ± 0.33	3.26 ± 0.56	3.21 ± 0.48	3.16 ± 0.54	3.16 ± 0.46
ThoraxML	2.68 ± 0.68	2.21 ± 0.74	2.04 ± 0.70	2.01 ± 0.62	1.94 ± 0.57

0.18, 0.31, 0.63 λ (Fig. 4); and a maximum MDC of 41 %, 45 %, 27 %, and 76 % correspondingly (Fig. 5).

3.3. LyE algorithm

Both the Rosenstein and Wolf algorithms exhibited poor to good relative reliability across all trial durations and gait variables, with ICCs ranging from 0.45 to 0.87 and 0.14 to 0.84, respectively (Fig. 3). For absolute reliability and measure sensitivity the Rosenstein and Wolf algorithms showed SEMs ranging from 0.02 to 0.16 and 0.09 to 0.63 λ respectively (Fig. 4); and MDCs ranging from 8 to 30 % and 19–76 % respectively (Fig. 5).

Grand mean LyE values across both sessions are provided in Table 2.

4. Discussion

The aim of this study was to assess how run duration, gait kinematic variable, and LyE algorithm selection influence the inter-session reliability and sensitivity of LDS in a young population during running. This study demonstrates that increasing run duration improved the inter-session reliability of LDS measures, with a minimum of 4 min needed to achieve moderate to good reliability regardless of gait variable or algorithm choice. However, the reliability of medio-lateral pelvis acceleration was good from 2 min onwards regardless of algorithm, highlighting how gait variable choice can influence reliability. Furthermore, algorithm choice affected absolute reliability and sensitivity, whereby applying the Rosenstein algorithm produced lower SEM and MDC when compared to the Wolf algorithm. Therefore, when assessing repeat LDS measures inter-session in young people during running, analysing a minimum of 4-min of data is suggested. In addition, selecting medio-lateral pelvis acceleration or KFA variables may allow for shorter run durations whilst retaining reliability and applying the Rosenstein algorithm could help to produce measures more sensitive to change.

This study identified that when assessing LDS in young people during running, longer durations are preferable to enhance the inter-session reliability. However, identifying the shortest duration with the minimum acceptable reliability could reduce the time and effort required to collect data and thus, be preferred in clinical use. Based on our results, the minimum run duration required to achieve moderate to good reliability regardless of gait variable and LyE algorithm choice is 4 min. This is represented by tighter 95 %CI from 4 min (280 strides) onwards (Fig. 3). When distinguishing between gait variables, KFAdom was not affected by duration and produced moderate to good reliability for all durations. KFAndom achieved this with trials of at least 2 min (138 strides) from which PelvisML consistently achieved good reliability. This supports previous studies which reported moderate to excellent inter-session reliability (ICC > 0.5) using durations of 95 s to 3 min (Hamacher et al., 2016; Nazary-Moghadam et al., 2020) and 100–170 strides (Fohrmann et al., 2022; Raffalt et al., 2018) when assessing knee angle and sacrum variables. However, in the current study only PelvisML analysed using the Rosenstein algorithm produced 95 %CI consistently above the moderate reliability baseline from 3 min onwards, suggesting this combination may be preferable to confidently achieve moderate reliability. It should be noted that although only the Rosenstein algorithm produced 95 %CI above the moderate baseline, distinguishable differences between algorithms were not apparent. This was due to Rosenstein ICC point estimates falling within the bounds of the Wolf 95 %CI. These results suggest that a minimum of 4 min is needed to achieve moderate reliability regardless of the gait variable or algorithm, yet a choice of gait variable can potentially reduce this minimum time requirement. For instance, moderate reliability was achieved using 2-min trial durations for KFA, and good reliability achieved using 3-min durations for medio-lateral pelvis acceleration which produced more favourable 95 % CIs when using the Rosenstein algorithm. Therefore, when selecting a minimum run duration to measure LDS in a young population, practitioners should be guided by their research question and account for other methodological choices.

This study found differences in reliability between gait variables, further emphasising the importance of informed decision making in the calculation of LDS in young people. For instance, this study found that 2–3 min (138–209 strides) using the Rosenstein algorithm produced good reliability (ICC = 0.73–0.82) for pelvis, but only moderate to good (ICC = 0.59–0.87) for KFA. In contrast, a similar investigation found that the reliability of KFA was excellent (ICC = 0.92) and medio-lateral sacrum displacement (pelvis) was good (ICC = 0.68) for 170 strides of walking with the Rosenstein algorithm (Raffalt et al., 2018). This disagreement in KFA reliability may be explained by the LDS differences in walking and running. Raffalt et al. (2018) assessed walking which produces less variable movement and leads to higher reliability than running (Ekizos et al., 2018; Jordan et al., 2009). This is reflected in the mean KFA LyE value (1.18 λ) reported by Raffalt et al. (2018), which is lower than the current study's range of 2.44–2.59 λ (Table 2). This suggests that sacrum or pelvis gait variables may be a more robust choice to achieve reliable results when assessing LDS during running. This premise is supported further by how much stronger the effect of run duration was on pelvis medio-lateral acceleration LDS than other gait variables. The ICC for pelvis medio-lateral acceleration, from 2 min using the Wolf algorithm and 3 min using the Rosenstein algorithm, were outside and above the 95 %CI at 1 min. This highlights how longer durations of pelvis data produce significantly more reliable results than shorter durations, whereas other gait variables may not. Therefore, medio-lateral pelvis acceleration may offer superior reliability over KFA or medio-lateral thorax acceleration, particularly with run durations greater than 3 min.

The findings of this study suggest the choice of LyE algorithm has no clear effect on the relative reliability of LDS. As shown in Fig. 3, 95 %CI overlapped between algorithms indicating no meaningful differences were observed when analysing KFA and PelvisML. Both algorithms produced moderate to good reliability with similar 95 %CI for durations of greater than 1-min. This contradicts previous literature that reported the superiority of the Rosenstein algorithm, which produced good and excellent reliability for pelvis and KFA variables, respectively, while the Wolf algorithm produced poor reliability for both (Raffalt et al., 2018). However, this disparity in results may be due to Raffalt et al.'s inclusion of participants with knee osteoarthritis, who demonstrate differences in LDS to healthy controls (Fallah Yakhdani et al., 2010; Mahmoudian et al., 2016). Therefore, the Wolf algorithm may not produce reliable results in populations with knee pathology, perhaps due to added gait instability creating noisier signals which the Rosenstein algorithm may be more robust to (Rosenstein et al., 1992). Therefore, future researchers should be guided by their research population when deciding which LyE algorithm to use. It is possible that the Rosenstein algorithm may be best suited for assessing KFA and pelvis LDS in those rehabilitating from knee injury, although further research is needed.

If LDS measures are to be used to assess changes in gait function (e.g. during rehabilitation from injury) then the ability to assess changes within an individual both reliably and sensitively is crucial. Therefore, acceptable absolute inter-session reliability (i.e. low SEM) and repeat measure sensitivity (i.e. low MDC), is desirable. In this study, the choice of LyE algorithm and duration was shown to affect the SEM and MDC but influenced gait variables differently. The SEM of KFA measures were not affected by algorithm or duration, shown by the consistent values across durations and between algorithms (Fig. 4). The MDC of KFA followed a similar trend except when using the Wolf algorithm and durations of 1-min, whereby slightly worse MDC values were seen (Fig. 5). Generally, MDC for all variables was only affected by duration when analysed using the Wolf algorithm highlighting how the Rosenstein algorithm may be better suited for shorter durations. The Rosenstein algorithm's superiority over the Wolf algorithm for short data sets has been contested in the literature (Bruijn et al., 2012; Cignetti et al., 2012a, 2012b), with no clear recommendation of an optimal choice. However, use with short data sets was the rationale for its original development (Rosenstein et al., 1992). Furthermore, this study showed that in young people during running, the SEM of trunk variables (PelvisML and ThoraxML) were heavily affected by algorithm choice, with the Rosenstein algorithm generally producing lower values than Wolf (Fig. 4). Further, ThoraxML showed notably poorer SEM and MDC when using the Wolf algorithm (Fig. 4 and 5). Therefore, when assessing repeat measures of LDS in a young individual during running the use of either KFA or PelvisML may be more reliable, and the Rosenstein algorithm is more preferable for shorter durations.

It would be useful for future clinical use of repeat LDS measures in young people to discern any 'real' change in LDS from an expected normal change i.e. any change in LDS that is not due to systematic error. The results of this investigation show the MDC for PelvisML and KFA variables with durations of 3 to 5-min ranged from 8 to 24 % for the Rosenstein algorithm and 20 to 31 % for the Wolf algorithm (Fig. 5). This indicates a conservative expected variation of 24 % and 31 % between repeat LDS measures using the Rosenstein and Wolf algorithm, respectively. There are a lack of MDC values to compare to in the literature, however, using reported SEM and mean LyE values from studies with similar methods (Raffalt et al., 2018; van Schooten et al., 2013), a MDC of 17 to 33 % was calculated which aligns with the variation reported in this study (24 %). Therefore, future researchers and clinicians assessing session-to-session differences in medio-lateral pelvis acceleration and KFA LDS during running in young people can use these values to determine whether a measured change reflects a 'real' change or is merely due to error.

4.1. Limitations

This study suggests higher absolute reliability of trunk LyE estimates when using the Rosenstein algorithm than the Wolf algorithm. However, it has been argued that the Rosenstein algorithm might underestimate the magnitude of LyE due to limitations in its methods (Cignetti et al., 2012a). Low LyE values have the potential of a floor effect which may restrict the extent to which estimates can vary from session to session, in turn increasing the reliability. The low LyE values reported for trunk variables when analysed by the Rosenstein algorithm (Table 2) may support this and could explain the high reliability seen in these variables. The Rosenstein algorithm may simply be less sensitive in detecting changes in LDS within an individual. An ultimate consensus on which algorithm is most reliable for measures of LDS has still not been reached and algorithm choice should be driven by the research question, methodology, and population being studied. Therefore, the current findings can still be directly applied to assessing change in LDS specifically during running within a young population.

Further, the methods used in this study may only be applicable to athletes that compete in sports that involve running. The exclusion of a participant whose data was identified as an outlier and who was the only athlete to participate in a sport which did not involve running may support this. Therefore, the results are not generalisable to athletes in non-running sports and further work should explore the reliability of LDS in other movements e.g. jumping and hopping.

This study used an average inter-session period of 7 days. Therefore, the results may not be applicable in situations requiring longer durations between measurements (e.g. rehabilitation from injury). However, a duration of 7 days helps eliminate any effect of irregular activity levels that may be seen over longer durations in healthy athletes due to seasons, training load, and competitions.

5. Conclusion

This study found that longer running durations enhanced the inter-session reliability of LDS in healthy active young people and provides a recommended run duration of 4 min. Yet, the choice of medio-lateral pelvis acceleration or KFA variables may allow shorter trials to be collected, and applying the Rosenstein algorithm could help to produce measures more sensitive to change. The findings of this study can be used by those examining change in LDS of young people during running to ensure more robust and reliable results.

Funding

This study was funded by a collaborative PhD scholarship awarded to Adam S. Kennerley by Sheffield Hallam University (UK) and La Trobe University (Australia).

CRediT authorship contribution statement

Adam S. Kennerley: Writing – review & editing, Writing – original draft, Visualization, Validation, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Marcus Dunn:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Kane Middleton:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Kate E. Webster:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Jonathan Wheat:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We acknowledge Mr. Ian Brookes for their technical assistance.

Data availability

Data will be made available on request.

References

Afonso De Oliveira, E., Andrade, A. O., & Vieira, M. F. (2019). *Linear and nonlinear measures of gait variability after anterior cruciate ligament reconstruction*. https://doi.org/10.1016/j.jelekin.2019.03.007

Allet, L., Kim, H., Ashton-Miller, J. A., & Richardson, J. K. (2012). Which lower limb frontal plane sensory and motor functions predict gait speed and efficiency on uneven surfaces in older persons with diabetic neuropathy? *PM&R, 4*(10), 726–733. https://doi.org/10.1016/j.pmrj.2012.05.002

Ardern, C. L., Taylor, N. F., Feller, J. A., & Webster, K. E. (2014). *Fifty-five per cent return to competitive sport following anterior cruciate ligament reconstruction surgery: An updated systematic review and meta-analysis including aspects of physical functioning and contextual factors*. https://doi.org/10.1136/bjsports-2013-093398

Armand, S., Sangeux, M., & Baker, R. (2014). Optimal markers' placement on the thorax for clinical gait analysis. *Gait & Posture, 39*(1), 147–153. https://doi.org/10.1016/j.gaitpost.2013.06.016

Atkinson, G., & Nevill, A. M. (1998). Statistical methods for assessing measurement error (reliability) in variables relevant to sports medicine. *Sports Medicine, 26*(4), 217–238. https://doi.org/10.2165/00007256-199826040-00002

Bartlett, R., Wheat, J., & Robins, M. (2007). Is movement variability important for sports biomechanists? *Sports Biomechanics, 6*(2), 224–243. https://doi.org/10.1080/14763140701322994

Bramah, C., Tawiah-Dodoo, J., Rhodes, S., Elliott, J. D., & Dos'Santos, T.. (2024). The sprint mechanics assessment score: A qualitative screening tool for the in-field assessment of Sprint running mechanics. *The American Journal of Sports Medicine, 52*(6), 1608–1616. https://doi.org/10.1177/03635465241235525

Bruijn, S. M., Meijer, O. G., Beek, P. J., & Van Dieën, J. H. (2013). Assessing the stability of human locomotion: A review of current measures. *Journal of the Royal Society Interface, 10*(83), 20120999. https://doi.org/10.1098/rsif.2012.0999

Bruijn, S. M., Meijer, O. G., Rispens, S. M., Daffertshofer, A., & Van Dieën, J. H. (2012). Letter to the editor: "Sensitivity of the Wolf's and Rosenstein's algorithms to evaluate local dynamic stability from small gait data sets.". *Annals of Biomedical Engineering, 40*(12), 2505–2506. https://doi.org/10.1007/s10439-012-0586-4

Buckthorpe, M., Tamisari, A., & Villa, F. D. (2020). A ten task-based progression in rehabilitation after ACL reconstruction: From post-surgery to return to play – A clinical commentary. *International Journal of Sports Physical Therapy, 15*(4), 611–623. https://doi.org/10.26603/ijspt20200611

Castiglia, S. F., Trabassi, D., Conte, C., Gioiosa, V., Sebastianelli, G., Abagnale, C., … Serrao, M. (2024). Local dynamic stability of trunk during gait is responsive to rehabilitation in subjects with primary degenerative cerebellar ataxia. *The Cerebellum*. https://doi.org/10.1007/s12311-024-01663-4

Cignetti, F., Decker, L. M., & Stergiou, N. (2012a). Sensitivity of the Wolf's and Rosenstein's algorithms to evaluate local dynamic stability from small gait data sets. *Annals of Biomedical Engineering, 40*(5), 1122–1130. https://doi.org/10.1007/s10439-011-0474-3

Cignetti, F., Decker, L. M., & Stergiou, N. (2012b). 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. *Annals of Biomedical Engineering, 40*(12), 2507–2509. https://doi.org/10.1007/s10439-012-0665-6

Conn, J. M., Annest, J. L., & Gilchrist, J. (2003). Sports and recreation related injury episodes in the US population, 1997–99. *Injury Prevention, 9*(2), 117–123. https://doi.org/10.1136/ip.9.2.117

Dingwell, J. B., & Marin, L. C. (2006). Kinematic variability and local dynamic stability of upper body motions when walking at different speeds. *Journal of Biomechanics, 39*(3), 444–452. https://doi.org/10.1016/j.jbiomech.2004.12.014

Ekizos, A., Santuz, A., Schroll, A., & Arampatzis, A. (2018). The maximum Lyapunov exponent during walking and running: Reliability assessment of different marker-sets. *Frontiers in Physiology, 9*(AUG). https://doi.org/10.3389/fphys.2018.01101

Fallah Yakhdani, H. R., Bafghi, H. A., Meijer, O. G., Bruijn, S. M., Dikkenberg, N. V. D., Stibbe, A. B., … Van Dieën, J. H. (2010). Stability and variability of knee kinematics during gait in knee osteoarthritis before and after replacement surgery. *Clinical biomechanics, 25*(3), 230–236. https://doi.org/10.1016/j.clinbiomech.2009.12.003

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fohrmann, D., Hamacher, D., Sanchez-Alvarado, A., Potthast, W., Mai, P., Willwacher, S., & Hollander, K. (2022). Reliability of running stability during treadmill and overground running. *Sensors, 23*(1), 347. https://doi.org/10.3390/s23010347

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. *Physical Review A, 33*(2), 1134–1140. https://doi.org/10.1103/PhysRevA.33.1134

Full, R. J. (2002). Quantifying dynamic stability and maneuverability in legged locomotion. *Integrative and Comparative Biology, 42*(1), 149–157. https://doi.org/10.1093/icb/42.1.149

Hamacher, D., Hamacher, D., Törpel, A., Krowicki, M., Herold, F., & Schega, L. (2016). The reliability of local dynamic stability in walking while texting and performing an arithmetical problem. *Gait & Posture, 44*, 200–203. https://doi.org/10.1016/j.gaitpost.2015.12.021

Hilfiker, R., Vaney, C., Gattlen, B., Meichtry, A., Deriaz, O., Lugon-Moulin, V., Anchisi-Bellwald, A.-M., Palaci, C., Foinant, D., & Terrier, P. (2013). Local dynamic stability as a responsive index for the evaluation of rehabilitation effect on fall risk in patients with multiple sclerosis: A longitudinal study. *BMC Research Notes, 6*(1), 260. https://doi.org/10.1186/1756-0500-6-260

Hunter, B., Karsten, B., Greenhalgh, A., Burnley, M., & Muniz-Pumares, D. (2023). The application of non-linear methods to quantify changes to movement dynamics during running: A scoping review. *Journal of Sports Sciences, 41*(5), 481–494. https://doi.org/10.1080/02640414.2023.2225014

Hussain, V. S., Frames, C. W., & Lockhart, T. E. (2021). Length of time-series gait data on Lyapunov exponent for fall risk detection. *International Journal of Prognostics and Health Management, 12*(4). https://doi.org/10.36001/ijphm.2021.v12i4.2917. Article 4.

Jordan, K., Challis, J. H., Cusumano, J. P., & Newell, K. M. (2009). Stability and the time-dependent structure of gait variability in walking and running. *Human Movement Science, 28*(1), 113–128. https://doi.org/10.1016/j.humov.2008.09.001

Jordan, K., Challis, J. H., & Newell, K. M. (2006). Long range correlations in the stride interval of running. *Gait and Posture, 24*(1), 120–125. https://doi.org/10.1016/j.gaitpost.2005.08.003

Kang, H. G., & Dingwell, J. B. (2006). Intra-session reliability of local dynamic stability of walking. *Gait & Posture, 24*(3), 386–390. https://doi.org/10.1016/j.gaitpost.2005.11.004

Kennel, M. B., & Isabelle, S. (1992). Method to distinguish possible chaos from colored noise and to determine embedding parameters. *Physical Review A, 46*(6), 3111–3118. https://doi.org/10.1103/PhysRevA.46.3111

Kisser, R., & Bauer, R. (2012). *The burden of sport injuries in the European Union*.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine, 15*(2), 155. https://doi.org/10.1016/J.JCM.2016.02.012

Mahmoudian, A., Bruijn, S. M., Yakhdani, H. R. F., Meijer, O. G., Verschueren, S. M. P., & Van Dieen, J. H. (2016). Phase-dependent changes in local dynamic stability during walking in elderly with and without knee osteoarthritis. *Journal of Biomechanics, 49*(1), 80–86. https://doi.org/10.1016/j.jbiomech.2015.11.018

Mehdizadeh, S., & Sanjari, M. A. (2017). Effect of noise and filtering on largest Lyapunov exponent of time series associated with human walking. *Journal of Biomechanics, 64*, 236–239. https://doi.org/10.1016/j.jbiomech.2017.09.009

van Melick, N., Meddeler, B. M., Hoogeboom, T. J., Nijhuis-van der Sanden, M. W. G., & van Cingel, R. E. H. (2017). How to determine leg dominance: The agreement between self-reported and observed performance in healthy adults. *PLoS One, 12*(12), Article e0189876. https://doi.org/10.1371/journal.pone.0189876

Moraiti, C. O., Stergiou, N., Vasiliadis, H. S., Motsis, E., & Georgoulis, A. D. (2010). Anterior cruciate ligament reconstruction results in alterations in gait variability. *Gait & Posture, 32*(2), 169–175.

Myer, G. D., Paterno, M. V., Ford, K. R., Quatman, C. E., & Hewett, T. E. (2006). Rehabilitation after anterior cruciate ligament reconstruction: Criteria-based progression through the return-to-sport phase. *The Journal of Orthopaedic and Sports Physical Therapy, 36*(6), 385–402. https://doi.org/10.2519/jospt.2006.2222

Nazary-Moghadam, S., Salavati, M., Esteki, A., Akhbari, B., Keyhani, S., & Zeinalzadeh, A. (2020). Reliability of knee flexion-extension lyapunov exponent in people with and without anterior cruciate ligament deficiency. *Journal of Sport Rehabilitation, 29*(2), 253–256. https://doi.org/10.1123/JSR.2018-0468

Raffalt, P. C., Alkjaer, T., Brynj, B., Jørgensen, L., Bartholdy, C., & Henriksen, M. (2018). *Day-to-day reliability of nonlinear methods to assess walking dynamics*. https://doi.org/10.1115/1.4041044

Raffalt, P. C., Senderling, B., & Stergiou, N. (2020). Filtering affects the calculation of the largest Lyapunov exponent. *Computers in Biology and Medicine, 122*, Article 103786. https://doi.org/10.1016/j.compbiomed.2020.103786

Reynard, F., & Terrier, P. (2014). Local dynamic stability of treadmill walking: Intrasession and week-to-week repeatability. *Journal of Biomechanics, 47*(1), 74–80. https://doi.org/10.1016/j.jbiomech.2013.10.011

Rosenstein, M. T., Collins, J. J., & De Luca, C. J. (1992). *A practical method for calculating largest Lyapunov exponents from small data sets*.

van Schooten, K. S., Rispens, S. M., Pijnappels, M., Daffertshofer, A., & van Dieen, J. H. (2013). Assessing gait stability: The influence of state space reconstruction on inter- and intra-day reliability of local dynamic stability during over-ground walking. *Journal of Biomechanics, 46*(1), 137–141. https://doi.org/10.1016/j.jbiomech.2012.10.032

Schrijvers, J. C., Van Den Noort, J. C., Van Der Esch, M., & Harlaar, J. (2021). Neuromechanical assessment of knee joint instability during perturbed gait in patients with knee osteoarthritis. *Journal of Biomechanics, 118*, Article 110325. https://doi.org/10.1016/j.jbiomech.2021.110325

Sharifi, M., & Shirazi-Adl, A. (2021). Knee flexion angle and muscle activations control the stability of an anterior cruciate ligament deficient joint in gait. *Journal of Biomechanics, 117*, Article 110258. https://doi.org/10.1016/j.jbiomech.2021.110258

Stergiou, N. (2016). *Nonlinear analysis for human movement variability*. CRC Press.

Stergiou, N., & Decker, L. M. (2011). Human movement variability, nonlinear dynamics, and pathology: Is there a connection? *Human Movement Science, 30*(5), 869–888. https://doi.org/10.1016/j.humov.2011.06.002

Strongman, C., & Morrison, A. (2020). A scoping review of non-linear analysis approaches measuring variability in gait due to lower body injury or dysfunction. *Human Movement Science, 69*, Article 102562. https://doi.org/10.1016/j.humov.2019.102562

Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision, 35*(17), 2503–2522.

UNO Biomechanics. (2023). UNO Biomechanics nonlinear analysis toolbox. MATLAB Central file exchange. https://www.mathworks.com/matlabcentral/fileexchange/71907-uno-biomechanics-nonlinear-analysis-toolbox.

Weir, J. P. (2005). Quantifying test-restest reliability using the intraclass correlation coefficient and the SEM. *The Journal of Strength and Conditioning Research, 19*(1), 231–240.

Whittaker, J. L., Woodhouse, L. J., Nettel-Aguirre, A., & Emery, C. A. (2015). Outcomes associated with early post-traumatic osteoarthritis and other negative health consequences 3–10 years following knee joint injury in youth sport. *Osteoarthritis and Cartilage, 23*(7), 1122–1129. https://doi.org/10.1016/j.joca.2015.02.021

Winter, L., Taylor, P., Bellenger, C., Grimshaw, P., & Crowther, R. G. (2023). The application of the Lyapunov exponent to analyse human performance: A systematic review. *Journal of Sports Sciences, 41*(22), 1994–2013. https://doi.org/10.1080/02640414.2024.2308441

Wolf, A., Swift, J., Swinney, H. L., Vastano, J., Wolft, A., Swift, J. B., & Vastano, J. A. (1985). Determining Lyapunov exponents from a time series. *Physica D: Nonlinear Phenomena, 16*(3). https://doi.org/10.1016/0167