



Assessing gait stability: The influence of state space reconstruction on inter- and intra-day reliability of local dynamic stability during over-ground walking

Kimberley S. van Schooten, Sietse M. Rispens, Mirjam Pijnappels*, Andreas Daffertshofer, Jaap H. van Dieën

^a MOVE Research Institute Amsterdam, Faculty of Human Movement Sciences, VU University Amsterdam, Amsterdam, The Netherlands

ARTICLE INFO

Article history:

Accepted 26 October 2012

Keywords:

Fall risk
Gait stability
Lyapunov exponents
Repeatability
Test–retest

ABSTRACT

Estimating local dynamic stability is considered a powerful approach to identify persons with balance impairments. Its validity has been studied extensively, and provides evidence that short-term local dynamic stability is related to balance impairments and the risk of falling. Thus far, however, this relation has only been proven on group level. For clinical use, differences on the individual level should also be detectable, requiring reliability to be high. In the current study, reliability of short-term local dynamic stability was investigated within and between days. Participants walked 500 m back and forth on a straight outdoor footpath, on 2 non-consecutive days, and 3D linear accelerations were measured using an accelerometer (DynaPort MiniMod). The state space was reconstructed using 4 common approaches, all based on delay embedding. Within-session intra-class correlation coefficients were good (≥ 0.70), however between-session intra-class correlation coefficients were poor to moderate (≤ 0.63) and influenced by the reconstruction method. The same holds for the smallest detectable difference, which ranged from 17% to 46% depending on the state space reconstruction method. The best within- and between-session intra-class correlation coefficients and smallest detectable differences were achieved with a state space reconstruction with a fixed time delay and number of embedding dimensions. Overall, due to the influence of biological variation and measurement error, the short-term local dynamic stability can only be used to detect substantial differences on the individual level.

© 2012 Elsevier Ltd. Open access under the [Elsevier OA license](http://creativecommons.org/licenses/by/3.0/).

1. Introduction

Estimating local dynamic stability is a promising approach to quantify someone's ability to withstand perturbations and avert falling (Dingwell and Cusomano, 2000). Local dynamic stability could thus be a useful measure to identify individuals with balance impairments or at risk of falling. It is considered to capture how the neuromuscular system responds instantaneously to small perturbations by the average exponential rate of separation after such perturbations in state space (Dingwell and Kang, 2007), often referred to as maximum Lyapunov exponent.

The validity of local dynamic stability during gait has been studied extensively in modeling (Bruijn et al., 2012; Kurz et al., 2010; Roos and Dingwell, 2010; Su and Dingwell, 2007), experimental (Chang et al., 2010; McAndrew et al., 2011; Sloat et al., 2011; van Schooten et al., 2011) and observational studies (Kang and Dingwell, 2008, 2009; Lockhart and Liu, 2008; Toebe et al., 2012).

These studies provided evidence that short-term, finite-time local dynamic stability (λ_s) is related to impairments of balance and the risk of falling, on the group level. However, clinical use, i.e. identifying individuals at risk or monitoring intervention effects, requires detection of individual changes, and thus high reliability. Only a single study (Kang and Dingwell, 2006) investigated test–retest reliability of λ_s of trunk kinematics, and reported intra-class correlation coefficients $ICC(s)$ between 0.45 and 0.85 for estimates obtained from 1 to 5 min walking. As $ICCs$ between 0.71 and 0.96 have been reported for other commonly used balance measures (Henriksen et al., 2004; Moe-Nilssen and Helbostad, 2005; Stolze et al., 1998), this seems promising. However, this study only assessed within-day reliability, whereas clinical practice often requires measurements on several days, which increases within-subject variation. In addition, test–retest reliability was investigated during treadmill walking, which might be less variable and more stable than over-ground walking (Dingwell et al., 2001).

An important aspect in estimating local dynamic stability is a reconstruction of the state space, a set of vectors describing every point in time uniquely. However, since various linear combinations of a system's state variables may be used to span

* Corresponding author. Tel.: +31 20 59 88 467.

E-mail address: m.pijnappels@vu.nl (M. Pijnappels).

its state space, different reconstruction methods may yield different solutions (Kantz and Schreiber, 2004) rendering their comparison a challenge. In fact, even linear transforms of the state space may affect the mean and standard error of the local dynamic stability estimate (Gates and Dingwell, 2009; Rosenstein et al., 1994). Nonetheless, there is no consensus in literature yet on how to reconstruct the state space for gait dynamics. We hence employed different, commonly used techniques to address our primary aim, namely to assess test–retest reliability of short-term, finite-time local dynamic stability based on trunk accelerations during over-ground gait, within and between sessions.

2. Methods

20 healthy young adults (aged 28.5 ± 3.3 years) participated in this study. They all signed informed consent and the protocol was approved by the local ethics committee.

Participants walked 500 m forth and back (2 trials) over a straight outdoors footpath at their preferred walking speed. Measurements were done on 2 non-consecutive days, with 2 to 30 weeks between measurement days, resulting in 2 sessions of 2 trials each. Participants fitted themselves with a portable wireless accelerometer (DynaPort MiniMod, McRoberts, Den Haag, the Netherlands) attached to a neoprene belt around their pelvis with the accelerometer over the spine at the level of L5. This accelerometer measured linear accelerations in 3D over a range of ± 6 g, and sampled at a rate of 100 samples/s. To avoid the effect of fatigue, participants were seated for minimally 2 min between the 2 trials of a session.

Data were analyzed using MATLAB (version 7.12, The MathWorks BV, Natick, USA). Heel strikes were determined as the maximal vertical acceleration of the trunk, and the middle 200 strides were analyzed. As differences in the number of samples is known to cause a bias in estimating local dynamic stability (Bruijn et al., 2009; England and Granata, 2007), strides were time-normalized to 99 samples on average per stride (i.e. to the mean stride length across subjects in this study).

To investigate the effect of state space reconstruction on test–retest reliability of λ_s , different state spaces were reconstructed using time delay methods (Rosenstein et al., 1994). To this end, we either started off with a 1D signal, e.g., acceleration in mediolateral direction, or the full 3D signal, i.e. acceleration in all 3 directions, and created their time delayed copies that were embedded in a high-dimensional space (Kantz and Schreiber, 2004; Rosenstein et al., 1993):

$$X(t) = [x(t), x(t+J), \dots, x(t+(m-1) \cdot J)]$$

where $X(t)$ is the state of the system at discrete time t , $x(t)$ is the original time series, J is the time delay, and m is the number of embedding dimensions. For the number of embedding dimensions, we either used the global false nearest neighbors method (GFNN) for every individual signal (Kennel et al., 1992), or fixed the dimensions across data (see below). Reconstruction further differed by the choice of the embedding delay that was either determined individually for every signal via the first minimum of its average mutual information (I_{min}) (Fraser and Swinney, 1986), or fixed to a certain value that agreed for all signals (see again below). In sum, test–retest reliability was investigated for 4 state space reconstruction methods:

- (1) accelerations in mediolateral direction, dimension determined by GFNN and time delay by I_{min} ;
- (2) accelerations in mediolateral direction, a fixed number of dimensions (median of GFNN for our data=7), and time delay determined by I_{min} ;
- (3) accelerations in mediolateral direction, 7 dimensions, with a fixed time delay (median of I_{min} for our data=6 samples), and
- (4) a state space using accelerations in all 3 directions, 9 dimensions (since one time delayed copy plus the original 3D data is lower than the required 7), and a fixed time delay ($\frac{1}{4}$ of the mean stride time=24 samples).

In order to estimate local dynamic stability, for each point in state space the nearest neighbor was found, and the Euclidian distance between these nearest neighbors was determined as a function of time. In the case of local instability nearest neighbors will diverge, in particular along the most unstable direction. It is the rate of this divergence that specifies the linear part of the dynamics under study, i.e. the divergence evolves exponentially:

$$d(t) \approx d_0 e^{\lambda t} \text{ or alternatively } \ln d(t) \approx \lambda t + \ln d_0$$

where $d(t)$ is the divergence between the neighboring points at time t and d_0 is the initial distance between the nearest neighbors. The local dynamic stability λ_s is thus the divergence curve's slope λ that can be determined via a conventional least-squares fit over a certain (initial) time span (Rosenstein et al., 1993).

More specifically, λ_s as used in gait is the slope over a span of 0–0.5 strides (Bruijn et al., 2009a). In general, a positive λ_s indicates that systems with initially small differences will soon behave quite differently, hence such systems are considered locally unstable (Rosenstein et al., 1993). By contrast, if λ_s is negative, orbits will converge in time and the system is considered locally stable. A schematic is given in Fig. 1.

To explore the effect of within and between day measurements on λ_s , a 2 (sessions) \times 2 (trials) repeated measures ANOVA was done separately for all 4 methods (PASW statistics, version 18.0, SPSS Inc., Chicago, USA). To quantify test–retest reliability ICCs (2,1) absolute agreement (McCraw and Wong, 1996) were calculated within and between sessions,

$$ICC(2,1) = \frac{MS_R - MS_E}{MS_R + (2-1)MS_E + \frac{2}{n}(MS_C - MS_E)}$$

where MS_R is the mean square for rows (i.e. within subjects), MS_E the mean square error, MS_C is the mean square for columns (i.e. between subjects), and n the number of subjects. ICCs quantify the resemblance between 2 measures and range from 0 to 1, where an ICC of 1 indicates identical outcomes on 2 repeated measurements. In addition, the smallest detectable difference (SDD) was estimated, i.e., the smallest individual change that can be determined with 95% confidence as

$$SDD = \pm 1.96 \cdot \sqrt{2} \cdot SEM$$

where SEM is the standard error of mean. For all statistical tests a p -value < 0.05 was considered significant.

3. Results

No significant differences in λ_s were found between sessions and trials. Means of λ_s ranged from 0.64 to 1.55 (Table 1). Within sessions, ICCs ranged from 0.74 to 0.92 (Table 2). Between sessions, ICCs were considerably lower, and ranged from 0.38 to 0.63 (Table 3). The SDD expressed as percentage of the mean λ_s ranged from 8% to 46% (Tables 2 and 3). The best SDD values were observed for the full 9D state space (within 12%, between 20%) and method 3 (9.5% and 20%, respectively).

4. Discussion

Local dynamic stability is considered a promising approach to quantify one's ability to withstand perturbations and avert falling. However, its between sessions reliability has not been studied yet, although it is important for the application of this method as a scientific and diagnostic tool. Hence, we assessed the test–retest reliability of short-term finite-time local dynamic stability, within and between sessions, for several commonly used state space reconstructions based on trunk accelerations during over-ground walking. The within-session ICCs found in this study were all above 0.70, indicating that test–retest reliability of short-term local dynamic stability within a session is good (Fleiss, 1986). The values are in agreement with Kang and Dingwell (2006), who found an ICC of ~ 0.75 during treadmill walking for a trial length comparable to the one used in this study (± 3 min). Between sessions, the reliability was lower ($ICCs \leq 0.63$), i.e. poor for methods 1, 2, and 3 and moderate for method 4 (Fleiss, 1986). This shows that, depending on the state space reconstruction, local dynamic stability can be measured reliably enough to assess differences on the group level.

On the individual level, the smallest individual change the measure is capable of measuring (SDD) ranged from 17% to 46%. These percentages are fairly high and, therefore, only substantial changes in an individual can be interpreted as meaningful using this measure. To put this individual change in perspective, the difference in λ_s between young and older adults is 35%–50% (Kang and Dingwell, 2008, 2009), but the difference between older fallers and non-fallers is smaller, between 6% and 20% (Lockhart and Liu, 2008; Toebes et al., 2012). Thus, local dynamic stability can only detect substantial changes on the individual level, which might not be realistic. This problem could possibly be solved by

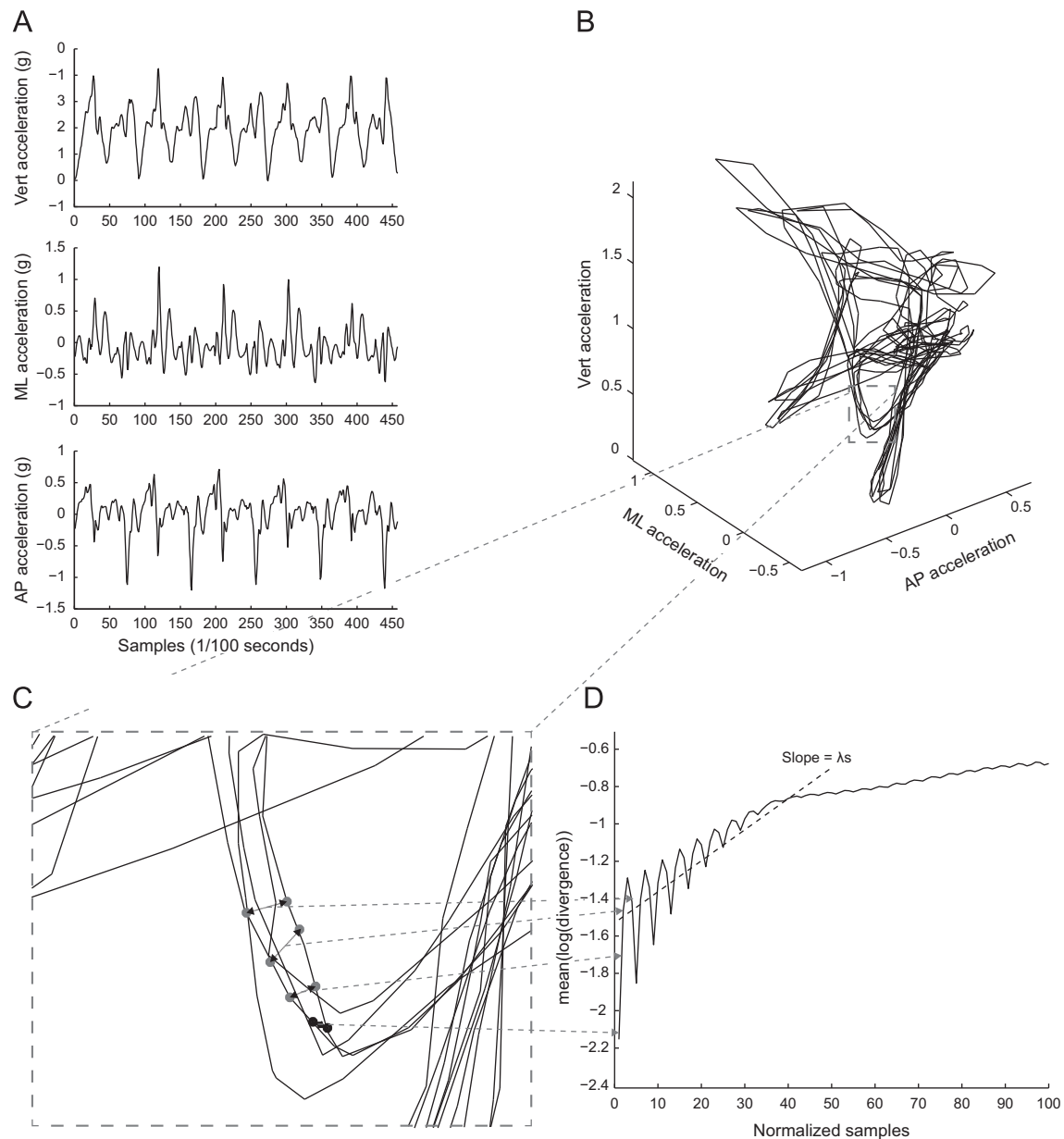


Fig. 1. Estimation of local dynamic stability. Original time series depicted in (A) are normalized to on average 99 samples per stride and embedded in 3D state space (B). In a local region (C), for each point in state space the nearest neighbor is found, and the mean logarithmic divergence is tracked over time (D). λ_s is calculated as the slope of the divergence curve between 0 and 0.5 strides. Vert=vertical direction, ML=mediolateral direction & AP=anteroposterior direction.

Table 1

Mean (and standard deviation) of λ_s per method, where state space reconstruction is based on (1) mediolateral accelerations with dimensions determined by $GFNN$ and time delay determined by I_{min} ; (2) mediolateral accelerations with 7 dimensions and time delay determined by I_{min} ; (3) mediolateral accelerations with 7 dimensions and a time delay of 6 samples, and (4) 3D accelerations, 9 dimensions and a time delay of 24 samples.

Method	Mean λ_s
1	1.53 (0.31)
2	1.55 (0.30)
3	1.37 (0.12)
4	0.64 (0.06)

measuring individuals on several days and averaging results, however such a method would come at a cost.

As expected, between sessions test–retest reliability was lower than within sessions and $SDDs$ were higher. One explanation

Table 2

Intra-class correlations (ICC) and smallest detectable differences (SDD) within sessions, where state space reconstruction is based on (1) mediolateral accelerations with dimensions determined by $GFNN$ and time delay determined by I_{min} ; (2) mediolateral accelerations with 7 dimensions and time delay determined by I_{min} ; (3) mediolateral accelerations with 7 dimensions and a time delay of 6 samples, and (4) 3D accelerations, 9 dimensions and a time delay of 24 samples.

Method	Session 1 (trial 1 vs. 2)				Session 2 (trial 3 vs. 4)			
	ICC	p	SDD	$\frac{SDD}{\text{Mean } \lambda_s}$	ICC	p	SDD	$\frac{SDD}{\text{Mean } \lambda_s}$
1	0.79	< 0.001	0.58	0.38	0.78	< 0.001	0.48	0.31
2	0.81	< 0.001	0.64	0.41	0.74	< 0.001	0.50	0.33
3	0.92	< 0.001	0.12	0.08	0.83	< 0.001	0.15	0.11
4	0.80	< 0.001	0.07	0.11	0.79	< 0.001	0.09	0.13

could be the reattachment of the accelerometer. To mimic realistic use, participants were asked to attach the accelerometer themselves. Reattachment of the accelerometer between sessions

Table 3

Intra-class correlations (ICC) and smallest detectable differences (SDD) between sessions, where state space reconstruction is based on (1) mediolateral accelerations with dimensions determined by GFNN and time delay determined by I_{min} ; (2) mediolateral accelerations with 7 dimensions and time delay determined by I_{min} ; (3) mediolateral accelerations with 7 dimensions and a time delay of 6 samples, and (4) 3D accelerations, 9 dimensions and a time delay of 24 samples.

Method	Trial 1 vs. 3				Trial 2 vs. 4			
	ICC	<i>p</i>	SDD	$\frac{SDD}{Mean \lambda_s}$	ICC	<i>p</i>	SDD	$\frac{SDD}{Mean \lambda_s}$
1	0.51	0.011	0.59	0.39	0.45	0.025	0.71	0.46
2	0.54	0.008	0.54	0.35	0.46	0.022	0.69	0.45
3	0.38	0.051	0.31	0.23	0.44	0.026	0.23	0.17
4	0.53	0.008	0.13	0.21	0.63	0.001	0.12	0.19

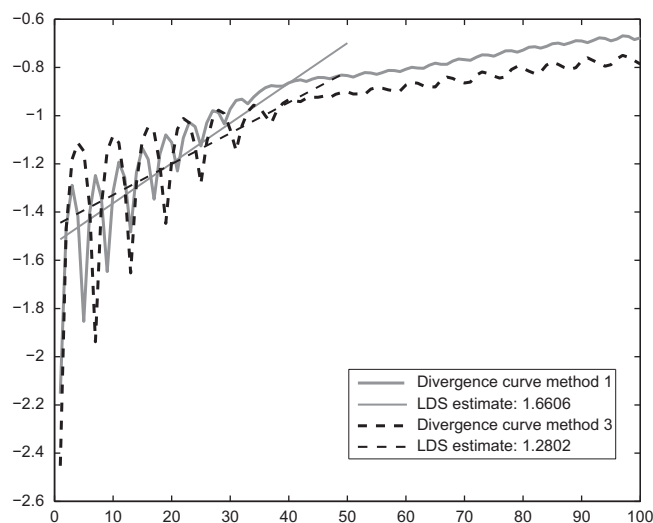


Fig. 2. Effect of state space reconstruction on negative peaks in the divergence curve. Two divergence curves of the same mediolateral acceleration signal. In gray, divergence calculated for a state space reconstruction with 9 dimensions and a time delay of 4 samples (following method 1), and in dashed black for a state space reconstruction with 7 dimensions and a time delay of 6 samples (as in method 3).

may have influenced measurement errors if the belt was secured more firmly or loosely to the body, in addition, differences in sensor orientation may have also caused systematic differences between measurements. The latter is supported by the higher ICCs and lower SDDs between sessions for the full 9D state space where all original 3D directions were used to estimate λ_s , which reduces the influence of sensor orientation. Although the attachment of the accelerometer is inherent to this method of measuring, our results call for consistency in (re-)attachment and/or awareness of this influence. Lower between sessions test–retest reliability could have also been caused by true biological variation, possibly related to the time of the day or even the individual's mood. More research is needed to distinguish all possible explanations.

Methods 3 and 4, which used fixed delays and embedding dimensions for state space reconstruction, yielded the best within- and between-session test–retest reliability, as well as the best smallest detectable differences. To further explore these differences between state space reconstruction methods, we assessed the divergence curves of the different methods. Fig. 2 depicts 2 divergence curves obtained from the same data but with a different time delay and embedding dimension (method 1 vs. method 3). In the first 50 samples (0–0.5 strides), there are marked negative peaks that appear at distinct intervals

(i.e. $[0,1,2,\dots,m] \cdot J + 1$, with m being the embedding dimensions and J the time delay). This could have been caused by noise-related errors that pertain due to the iterative nature of embedding and may have yielded a (spurious) dependency of λ_s on the selected time delay and number of embedding dimensions, leading to variance in outcomes between measurements when using different time delays and amount of dimensions. Unfortunately, averaging divergence over multiple nearest neighbors (Kantz and Schreiber, 2004) did not resolve this issue, and therefore it appears advisable to use a state space with a fixed number of dimensions and time delay.

In the current study we only used accelerations for reconstruction of the state space as accelerometry is readily applicable in large-scale studies and clinical practice. As signal characteristics are different for, e.g., angular velocity, reliability might improve when adding these variables for estimation of λ_s . Future research should address this, what we believe, important issue. λ_s in gait analysis is an approximation of the ‘true’ maximal Lyapunov exponent (Bruijn et al., 2012; Dingwell and Cusomano, 2000; McAndrew et al., 2011; Rosenstein et al., 1993), and this approximation might be susceptible to noise (Bruijn et al., 2012). Therefore, new methods for estimating local dynamic stability from gait data need further investigation. This may also increase reliability. We note that several studies have used kinematic variability together with short-term local dynamic stability to quantify stability (Toebes et al., 2012; van Schooten et al., 2011). This will probably influence test–retest reliability but as reliability of variability is fairly low (Brach et al., 2008), its beneficial capacities are yet unclear.

Overall, our results show that the estimation of local dynamic stability deserves improvement but nonetheless can be used to detect group differences. On the individual level, only large differences can be detected, making this measure not yet suited for the evaluation of interventions on the individual level or use as a screening tool, due to the influence of biological variation and/or measurement error. In addition, state space reconstruction strongly influences test–retest reliability due to the appearance of negative peaks in divergence curves. Therefore, test–retest reliability is better when the state space is reconstructed with a fixed number of embedding dimensions and a fixed time delay.

Conflict of interest statement

None of the authors of this paper had any conflict of interest that could inappropriately influence (i.e., bias) the presented work.

Acknowledgments

Kim van Schooten, Sietse Rispens and Mirjam Pijnappels, were financially supported by a TOP-NIG grant (#91209021) from the Dutch Organization for Scientific Research (NWO). Andreas Daffertshofer was also supported by an NWO grant (#400-08-127).

References

- Brach, J.S., Perera, S., Studenski, S., Newman, A.B., 2008. The reliability and validity of measures of gait variability in community-dwelling older adults. *Archives of Physical Medicine and Rehabilitation* 89, 2293–2296.
- Bruijn, S.M., Bregman, D.J., Meijer, O.G., Beek, P.J., van Dieën, J.H., 2012. Maximum Lyapunov exponents as predictors of global gait stability: a modelling approach. *Medical Engineering & Physics* 34, 428–436.
- Bruijn, S.M., van Dieën, J.H., Meijer, O.G., Beek, P.J., 2009. Statistical precision and sensitivity of measures of dynamic gait stability. *Journal of Neuroscience Methods* 178, 327–333.
- Chang, M.D., Sejdic, E., Wright, V., Chau, T., 2010. Measures of dynamic stability: detecting differences between walking overground and on a compliant surface. *Human Movement Science* 29, 977–986.

- Dingwell, J.B., Cusomano, J.P., 2000. Nonlinear time series analysis of normal and pathological human walking. *Chaos* (Woodbury, NY) 10, 848–863.
- Dingwell, J.B., Kang, H.G., 2007. Differences between local and orbital dynamic stability during human walking. *Journal of Biomechanical Engineering* 129, 586–593.
- Dingwell, J.B., Cusumano, J.P., Cavanagh, P.R., Sternad, D., 2001. Local dynamic stability versus kinematic variability of continuous overground and treadmill walking. *Journal of Biomechanical Engineering* 123, 27–32.
- England, S.A., Granata, K.P., 2007. The influence of gait speed on local dynamic stability of walking. *Gait & Posture* 25, 172–178.
- Fleiss, J.L., 1986. *The Design and Analysis of Clinical Experiments*. John Wiley & Sons Inc, New York.
- Fraser, A.M., Swinney, H.L., 1986. Independent coordinates for strange attractors from mutual information. *Physical Review A* 33, 1134.
- Gates, D.H., Dingwell, J.B., 2009. Comparison of different state space definitions for local dynamic stability analyses. *Journal of Biomechanics* 42, 1345–1349.
- Henriksen, M., Lund, H., Moe-Nilssen, R., Bliddal, H., Danneskiold-Samsøe, B., 2004. Test–retest reliability of trunk accelerometric gait analysis. *Gait & Posture* 19, 288–297.
- Kang, H.G., Dingwell, J.B., 2006. Intra-session reliability of local dynamic stability of walking. *Gait & Posture* 24, 386–390.
- Kang, H.G., Dingwell, J.B., 2008. Effects of walking speed, strength and range of motion on gait stability in healthy older adults. *Journal of Biomechanics* 41, 2899–2905.
- Kang, H.G., Dingwell, J.B., 2009. Dynamic stability of superior vs. inferior segments during walking in young and older adults. *Gait & Posture* 30, 260–263.
- Kantz, H., Schreiber, T., 2004. *Nonlinear Time Series Analysis*. Cambridge Univ Pr, Cambridge.
- Kennel, M.B., Brown, R., Abarbanel, H.D.I., 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A* 45, 3403.
- Kurz, M.J., Markopoulou, K., Stergiou, N., 2010. Attractor divergence as a metric for assessing walking balance. *Nonlinear Dynamics, Psychology, and Life Sciences* 14, 151–164.
- Lockhart, T.E., Liu, J., 2008. Differentiating fall-prone and healthy adults using local dynamic stability. *Ergonomics* 51, 1860–1872.
- McAndrew, P.M., Wilken, J.M., Dingwell, J.B., 2011. Dynamic stability of human walking in visually and mechanically destabilizing environments. *Journal of Biomechanics* 44, 644–649.
- McGraw, K.O., Wong, S., 1996. Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1, 30.
- Moe-Nilssen, R., Helbostad, J.L., 2005. Interstride trunk acceleration variability but not step width variability can differentiate between fit and frail older adults. *Gait & Posture* 21, 164–170.
- Roos, P.E., Dingwell, J.B., 2010. Influence of simulated neuromuscular noise on movement variability and fall risk in a 3D dynamic walking model. *Journal of Biomechanics* 43, 2929–2935.
- Rosenstein, M.T., Collins, J.J., De Luca, C.J., 1993. A practical method for calculating largest Lyapunov exponents from small data sets. *Physica D: Nonlinear Phenomena* 65, 117–134.
- Rosenstein, M.T., Collins, J.J., De Luca, C.J., 1994. Reconstruction expansion as a geometry-based framework for choosing proper delay times. *Physica D: Nonlinear Phenomena* 73, 82–98.
- Sloot, L.H., van Schooten, K.S., Bruijn, S.M., Kingma, H., Pijnappels, M., van Dieën, J.H., 2011. Sensitivity of local dynamic stability of over-ground walking to balance impairment due to galvanic vestibular stimulation. *Annals of Biomedical Engineering* 39, 1563–1569.
- Stolze, H., Kuhtz-Buschbeck, J., Mondwurf, C., Jöhnk, K., Friege, L., 1998. Retest reliability of spatiotemporal gait parameters in children and adults. *Gait & Posture* 7, 125–130.
- Su, J.L.-S., Dingwell, J.B., 2007. Dynamic stability of passive dynamic walking on an irregular surface. *Journal of Biomechanical Engineering* 129, 802–810.
- Toebe, M., Hoozemans, M., Furrer, R., Dekker, J., van Dieën, J., 2012. Local dynamic stability and variability of gait are associated with fall history in elderly subjects. *Gait & Posture* 36, 527–531.
- van Schooten, K.S., Sloot, L.H., Bruijn, S.M., Kingma, H., Meijer, O.G., Pijnappels, M., van Dieën, J.H., 2011. Sensitivity of trunk variability and stability measures to balance impairments induced by galvanic vestibular stimulation during gait. *Gait & Posture* 33, 656–660.