Title: Variability of within-step acceleration and daily wellness monitoring in Collegiate American Football

Submission Type: Original Investigation

Authors: Murray, Andrew\textsuperscript{1,2}; Buttfield, Alec\textsuperscript{3}; Simpkin, A\textsuperscript{4}; Sproule, John\textsuperscript{1}; Turner, Anthony P\textsuperscript{1}

\textsuperscript{1}Institute of Sport, PE & Health Sciences, University of Edinburgh, Edinburgh, UK

\textsuperscript{2}University of Oregon, Athletics, 2727 Leo Harris Parkway, Eugene, Oregon, USA 97401

\textsuperscript{3}Institute of Sport, Exercise and Active Living College of Sport and Exercise Science, Victoria University

\textsuperscript{4}School of Mathematics, Statistics and Applied Mathematics, National University of Ireland, Galway, Ireland

Contact details:
Andrew Murray
University of Oregon
Athletics Department
2727 Leo Harris Parkway
Eugene
OR 97401
amurray2@uoregon.edu

Running Title: Step Variability & Wellness in American Football

Declarations of interest: None

Word Count: 3496

Abstract word Count: 192

Figures: 1
Tables: 2
Abstract

Objectives: It is commonplace to consider accelerometer load and any resultant neuromuscular fatigue in training programs. With these data becoming accepted in sport alongside wellness questionnaires this study aimed to investigate if a deeper analysis of the accelerometry data can provide actionable insight into training-induced disruptions.

Design: Accelerometer data from Collegiate American Football athletes (n=63) were collected during training and matches across a regular season.

Methods: These data were processed to: identify instances of high speed running, extract step waveforms from those sections, and determine the variability of those waveforms via a within- and between-section co-efficient of multiple determination. Athletes completed wellness questionnaires prior to sessions that were used to flag areas of muscle soreness as well as fatigue, or disturbed sleep quality. Linear mixed models were used to assess associations between intra stride variability and flags in wellness/soreness markers.

Results: An increase in acute (7d) load saw an increased stride variability in these athletes. Feeling less fatigued and/or lower muscle soreness was associated with higher stride variability.

Conclusion: The assessment of variability has the potential to identify athletes who are displaying physical symptoms that would indicate the need to modify training.
Introduction

Movement variability exists even in highly trained skills performed by elite athletes,\(^1\) this would suggest that gait would also reflect the theoretical principle of a ‘healthy’ amount of movement variability. Indeed, individuals with patellofemoral knee pain have been shown to exhibit reduced movement variability compared to a healthy group.\(^2\) Although subsequent studies have produced contradictory findings,\(^3\) movement variability has been shown to increase in subjects affected by patellofemoral pain when their pain is reduced through a therapeutic intervention\(^3\), suggesting there is an individual level of movement variability in gait and that variability is decreased when pain is present. Increased fatigue has been shown to lead to increased variability in knee kinematics during a cutting maneuver, which in turn will lead to a reduced ability to produce a controlled movement.\(^4\) Consequently, the use of movement variability as a clinical tool to identify when an individual has a less than optimal movement pattern, is entirely possible as long as the chosen measurement tool has sufficient resolution to identify significant changes in an individual’s movement variability.

Wireless accelerometry is a popular approach to continuously assess both proximal (e.g. trunk) and distal (e.g. tibial) mechanics in human locomotion unobtrusively. This approach is common in inertial measurement units that are used with athletes and are worn on the torso – typically incorporating accelerometers, global positioning systems (GPS), magnetometers, and gyroscopes. Using this approach the magnitude of peak accelerations have been validated\(^5\) which demonstrates that filtered data collected by a MinimaxX S4 unit (Catapult Sports, Australia) provides an acceptable means of assessing peak accelerations (CV=8.9%). An alternate unit (SPI HPU, GPSports, Canberra, Australia) has been shown to accurately identify temporal stride characteristics (contact time \(r=0.98\); flight time \(r=0.68\)) when compared to an instrumented treadmill\(^6\). Ankle movement was constrained through taping and two of the three variables examined (contact time and vertical stiffness) correctly identified side-to-side differences in stride.
characteristics. These findings confirm the ability of this and similar units incorporating GPS and accelerometers, to identify small but practically important differences in stride characteristics due to physical constraints within a laboratory setting. This is particularly useful to applied practitioners given the practical and economic aspects of accelerometer technology.

The coefficient of multiple determination (CMD) and related coefficient of multiple correlation (CMC) have previously been used to analyze many forms of cyclic kinematic and kinetic data that have ranged from an analysis of kinematic variability in gymnastics\(^7\) to electromyographic, kinematic and kinetic measures of ice hockey skating\(^8\). Assessing the variability of waveforms has previously been done with gait data and been shown to be valid as a measure of stride characteristics via a single tri-axial accelerometer mounted on the upper torso\(^9\). Such analysis examines the waveform in its entirety rather than at specific points such as at foot strike or toe-off, and therefore accurate identification of specific points within the gait cycle will be less influential on the result of the analysis. In addition, using CMD to determine waveform variability does not require the waveforms to be from a continuous time period. This is a crucial consideration when analyzing data collected in gameplay and training rather than controlled laboratory settings.

It is common place to take accelerometer and GPS-derived running loads into consideration for the management of athletes\(^{10,11}\). With these data becoming commonplace in the sporting world alongside wellness questionnaires\(^12\) and athletes self-reporting muscle symptoms. This study aimed to investigate if a deeper analysis of the accelerometry data can be used to explore relationships between load, wellness, soreness and stride variability to provide actionable insight into training induced disruptions.

**Methods**

*Participants*
Data from 63 American Football athletes (20.6±1.5 yrs; 102.4±20.1 kg; 186±7.7 cm) operating at the Division 1 level in the NCAA were collected across a regular season. Athletes provided informed consent to participate in data collection throughout the season as part of the athlete support process and the institutional ethics committee provided ethical approval for the research.

**Design**

Inertial measurement units (IMU) containing GPS and accelerometers (Optimeye S5, Catapult Sports, Australia) were worn for every field session. The data collected and used in these studies were from the tri-axial accelerometer (measured at 100 Hz). For the purposes of this observational study, with repeated measures on the participants, only data from the main training sessions (Tuesday and Wednesday) and the match (Saturday) were recorded. This means light walk-through sessions on Sunday and Thursday were excluded, as were Friday sessions that were short and light in comparison to other sessions.

**Methodology**

**Accelerometry**

The 100Hz accelerometer data were processed with a novel analysis tool developed specifically for identifying instances of high speed running and determining the variability of the remaining waveforms via a within-section and between-section CMD. The raw files were exported via the manufacturer’s software (Catapult Sports, Openfield software, version 1.11.1). A step frequency of 2.75 steps per second for at least five seconds was chosen as the lower limit for high speed running. This step frequency was chosen after pilot testing (with the aim of achieving a similar number of step waveforms available for further analysis as was achieved in previous applications of the analysis tool) and is in general agreement with previous research.
Accelerometer data from those sections of high speed running were analyzed to identify steps through identifying foot strike events via peaks in the vertical accelerometer data. The step waveforms likely to have been influenced by gameplay demands were identified as steps where the mean vertical acceleration in the first 20% of the step was at least 2 standard deviations greater or less than the mean vertical acceleration for the first 20% of all steps on that day – these were eliminated from the analysis. Step waveforms were separated into left and right-side steps by examining the lateral accelerations, with steps displaying a negative to positive acceleration around foot strike being designated right side steps and vice versa. The CMD was then calculated on the set of vertical (z-axis) step waveforms to determine the variability of those waveforms as per Kadaba and colleagues\textsuperscript{14}. CMD values were calculated for each session for each player. They were combined over sections of high speed running during each game and CMD values were calculated from the within and between-stride variability, and then averaged over all sections of high speed running and turned into percentage of variation to improve interpretability. The data were therefore hierarchical in nature, with strides nested within sections within games within players. However, section-level data were unavailable for analysis.

Different calculations of variability were performed, one to examine the waveform variability within each section of high speed running, another to examine the variability between sections of high speed running. In all calculations, higher CMD scores indicate less waveform variability. All calculations occurred on the vertical axis as it has been shown previously that this is the most sensitive as a load indicator\textsuperscript{9}.

\textit{Wellness}

Over the course of the season the athletes completed a wellness questionnaire on training days, as used previously in the literature\textsuperscript{12}. This recorded any areas of soreness as well as noting their fatigue, sleep quality and overall muscle soreness (1=poor, 5=good). As part of the wellness
questionnaire athletes noted any specific locations of soreness and then rated these in term of severity (1-10). Any area greater than a 5 out of 10 for pain triggered a ‘flag’ to the practitioners working with the athletes. These flags are considered compromised training days in this study.

Load

IMU determined daily workloads (Playerload™) were calculated and expressed as arbitrary units (AU) via the manufacturer’s software (Catapult Sports, Openfield software, version 1.11.1) for every session. Participants wore the same device during every training session and match. Rolling loads for acute and chronic periods were calculated before sub setting the data to the main training sessions and games. The acute period was defined as 7 days and the chronic as 21 in line with previous American Football research. 15

Statistical Analysis

All analyses were carried out using R v3.5 (R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria URL https://www.R-project.org). Since repeated measures per player were available, linear mixed models were used to account separately for within-player and between-player variability in CMD values, while investigating their association with wellness (fatigue, sleep, soreness) and load (acute (7-day average), chronic (21-day average) and acute-chronic workload ratio) on that day. A random intercept term for player was used to allow for different average CMD values between athletes, while random slope terms allowed for different changes over time in CMD between players. A random effect for side of measurement was tested but led to convergence issues, hence it was included only as a fixed effect. To account for nonlinear changes in CMD over time, quadratic and cubic time terms were included as fixed and random effects. An AR1 process was included for within-subject variability to account for auto-regressive aspect of CMD during the
period of measurement. In each model, we also controlled for the number of strides, number of sections and side of measurement (left/right leg) to account for confounding. The association between each measure of wellness and load with CMD are presented as coefficients with 95% confidence intervals and p-values. Model residuals were checked to validate the assumptions underlying the linear mixed model. In order to compare between the load and wellness markers, we took the z-score of each of these (fatigue, sleep, soreness, ACWR, acute load, chronic load) and repeated analysis, with the resulting coefficients plotted showing the effect of a 1 standard deviation change in exposure. A secondary analysis focused on compromised training. Here a generalized linear mixed model was used to model flagged injury status against unflagged status for hamstring, ankle and foot injuries separately. The key variables under examination were within and between stride CMD measured during the flagged and unflagged strides. In each model, we included day of measurement as a fixed effect and used a random effect for athlete to allow each to have their own intercept. A logit link was used to model the three (hamstring, ankle and foot) binary outcomes (injured v not injured), and odds ratios are reported alongside 95% confidence intervals and p-values. Model residuals were again checked to validate the assumptions underlying the mixed model.

**Results**

Descriptive statistics for the key variables in the study are given in Table 1. All wellness variables had a mean of ~3, while acute (7-day) load was slightly higher than chronic (21-day) load. There were 4.94±5.75 (±SD) sections of high speed running per session across players on average, with a mean of 47.65±69.68 strides within a section. Figure 1 shows the nonlinear changes in CMD over the period of measurement, with similar patterns of change for within-stride and between-stride CMD.

***TABLE 1 NEAR HERE***

***FIGURE 1 NEAR HERE***
Wellness

There was some evidence for an inverse relationship between fatigue and between-stride CMD. A one-point increase in fatigue score (i.e. feeling better) being related to a 0.508% decrease in between-stride CMD (increased variability; table 2; 95% CI -0.953, -0.063%, p=0.025). There was no evidence for a relationship between sleep score and either within- or between-stride CMD. Finally, there was evidence for a negative association between soreness and CMD. A one-point increase in soreness score (i.e. less sore) was related to a 0.337% decrease in mean within-stride CMD (increased variability; table 2; 95% CI -0.670, -0.005%, p=0.047) and a 0.356% decrease in mean between stride CMD (table 2; 95% CI -0.752, 0.039%, p=0.078).

***TABLE 2 NEAR HERE***

ACWR had a negative effect on both within and between stride CMD, with a 1 unit increase in ACWR associated with a 6.849% decrease in mean within-stride CMD (increased variability; 95% CI -8.580, -5.117%, p<0.001) and a 7.257% decrease in mean between stride CMD (increased variability; 95% CI -9.355, -5.160%, p<0.001). Acute load (7-day average) was also associated with within- and between stride variability. A one unit increase in acute load was related to a 0.012% decrease in mean within-stride CMD (increased variability; 95% CI -0.016, -0.009%, p<0.001) and a 0.013% decrease in mean between stride CMD (increased variability; 95% CI -0.017, -0.010%, p<0.001). Finally, an increase in chronic load (21-day average) was also inversely related to within- and between-stride CMD. A one unit increase in chronic was associated with a 0.007% decrease in mean within-stride CMD (increased variability; 95% CI -0.011, -0.002%, p=0.002) and a 0.005% decrease in mean between stride CMD (increased variability; 95% CI -0.011, 0.000%, p=0.034).

*** FIGURE 2 NEAR HERE***
Comparing load and wellness

In order to compare the standardized coefficients across load and wellness where each exposure variable has been z-scored with the resulting coefficients showing the effect of a 1 standard deviation change in exposure (Figure 2). In this plot, the coefficients can be better compared. From Figure 2, acute load appears to have the strongest effect on within-stride CMD (-1.400%, 95% CI -1.751, -1.048%; p<0.001), followed by ACWR (-1.055%, 95% CI -1.322, -0.788%; p<0.001) and chronic load (-0.659%, 95% CI -1.079, -0.239%; p=0.002), with wellness measures having a weaker (per-SD) effect on CMD. Similarly, for between-stride CMD, load had a stronger effect in the same order, with acute being strongest (-1.493%, 95% CI -1.914, -1.073%; p<0.001) followed by ACWR (-1.118%, 95% CI -1.441, -0.795%; p<0.001) and chronic load (-0.532%, 95% CI -1.022, -0.041%; p=0.034).

Compromised training

Table 3 summarizes the models of compromised training and the effect of within and between stride CMD on these episodes. There were 9, 22 and 26 flagged hamstring, ankle, and foot injuries respectively. There was no strong evidence for an association between within or between stride CMD on any of the injury sites. However, given the small number of episodes, this analysis is underpowered. Within the sample, a one unit increase in between stride CMD was related to 3 times the odds of compromised training (odds ratio 3.111), but the interval estimate here is extremely wide (95% CI 0.297, 32.553) due to so few (n=9) hamstring episodes.

Discussion

The purpose of this study was to determine if analysis of the accelerometry data can provide actionable insight into training induced disruptions with no further testing on the athlete. This study has presented novel data showing that variability in stride detected by commonly used
accelerometers is associated with fatigue, soreness and training load. The ability to identify times when an athlete is at risk of injury or requires a training modification to maximize their performance in subsequent activities (whether that be a reduction or increase to their training load) is crucial in the preparation of athletes for competition.

**Load & Wellness**

The more fatigued athletes reported being the lower their stride variability. Previously with fatigue it has been shown that along with increased leg stiffness, the vertical motion of the CoM significantly reduces with prolonged exhaustive running.\(^\text{16}\) However, few studies have previously used trunk accelerometry to assess running related fatigue.\(^\text{17-19}\) In contrast to the current study, one study found a decrease in regularity of vertical CoM accelerations, when sub-elite distance runners underwent a short but highly intensive track run to exhaustion.\(^\text{19}\) Similarly, another showed that treadmill running-induced fatigue results in anteroposterior trunk accelerations that are less regular from step-to-step and are less predictable.\(^\text{18}\) The final study showed that CoM movement could accurately estimate increases in metabolic work during an incremental running protocol to exhaustion.\(^\text{17}\) It may be that the increased variability seen with these American Football players may signal a re-organization of motor strategies for the purpose of preserving performance (i.e. this increased stride variability may manifest as decreased variability in the upper body).

Previous research has demonstrated that fatigue alters the way player load is accumulated in Australian Rules Football matches.\(^\text{20}\) Other authors found that a one unit decrease in wellness Z-score resulted in a 4.9% (standard error 3.1%) and 8.6% (standard error 3.9%) decrease in player load and player load slow (running activity < 2 m.s\(^{-1}\)), respectively.\(^\text{21}\) Players with reduced wellness may maintain the running variables that they deem critical to performance but modify other aspects of activity profile such as change of speed, low speed running and/or body contact that were not measured in this study.\(^\text{22}\)
Within American Football specifically it has been shown that a one unit increase in wellness z-score and energy were associated with a trivial 2.3% and 2.6% increase in player load. A one unit increase in muscle soreness (players felt less sore) corresponded to a trivial 4.4% decrease in s-RPE training load. In addition, significant (p<0.05) differences in movement variables were demonstrated for individuals who responded more or less favorably on their rating of perceived wellness. In the current study while, there were no associations with sleep a decreased soreness resulted in an increase in variability – further investigations may look at the relationship between variability and sRPE directly.

An increase in load (both acute (7d) and chronic (21d) saw an increased variability in these team sport athletes. Although the mechanism underlying this increase in variability is currently unclear, it is roughly in agreement with previous theories, that suggest that a shift away from an individual’s optimal level of variability is indicative of a pathological state. A shift to an increased level of variability could be a sign of a noisy and irregular system, which has been demonstrated to be a characteristic of individuals who had undergone knee reconstructions to repair a damaged anterior cruciate ligament (possibly due to not being able to restore the proprioceptive pathways found in a healthy knee).

There is a high practical value to these findings as while current metrics do have the ability to predict injury risk, especially when examining cumulative load measures, they require a full training history to identify periods of load, (be that acute or chronic in nature), whereas if there is data missing or unavailable (such as when athletes are recruited into a squad on an intermittent basis or miss days through modified training) then the methods outlined here will still be able to identify individual athletes who have an elevated period of load compared to their normal training load (provided a baseline level of healthy movement variability has already been established).
Compromised Training

While there was an increased odds ratio of decreased variability in the presence of a flagged hamstring the analysis was too underpowered to draw a conclusion. Reduced variability would be expected for an acute injury. It has been observed that ACL deficient patients\textsuperscript{25} have less step-to-step variability in walking gait, inferring that they are being more “careful” when they were walking, trying to eliminate extraneous movements. The authors speculate that participants may be attempting to constrain movements and reduce step-to-step variability within the current results. The hamstring conditions likely indicate a compromised system. Further study may reveal if these flags are more indicative of chronic rather than acute conditions and so athletes have developed strategies to cope in these circumstances.

Limitations

The current investigation was limited to a single team over a single season, but still includes a total of 127,715 strides collected across 1177 sessions and 443 matches. A wider group would allow comparisons of differing training styles and approaches. Analyzing the occurrence of self-reported flags set at an arbitrary level (5/10) can be criticized as not everyone views discomfort in the same way and so potentially looking at an individual comparison may improve this metric.

Also, there were limited flags compared to the number of injuries that occur in collegiate football. The typical injury rates would suggest that 20\% of injuries are in the knee\textsuperscript{27} but these may be catastrophic one-off issues (i.e. ACL) rather than a degenerative issue that can be detected by flagging in a routine questionnaire. So, while early detection of issues as this study has shown possible is key, the differing positional demands and subsequent injury rates may need future studies to delineate the effects for particular positions in American Football in the context of injury history.
Practical Applications

The difference in the measures outlined is that predictions can be made from physical symptoms, but these track well with at least some of the subjective markers that athletes are giving. What is not known is how many athletes are not accurately flagging symptoms of soreness and so are going undetected in this analysis. In the absence of 100% disclosure from athletes the assessment of variability therefore has the potential to identify athletes who are displaying physical symptoms that would indicate the need to modify training. Conversely, it may be able to identify athletes who do satisfy flagging criteria but are showing no physical symptoms who therefore may not need training modifications.

Conclusions

This study has shown that stride variability is associated with fatigue and 7-day training load. Combining both objective and subjective methods is likely to enhance the predictive ability and become a very powerful tool within elite sport environments, and while further investigations into this are warranted, the assessment of variability has the potential to identify athletes who are displaying physical symptoms that would indicate the need to modify training.


Step Variability & Wellness in American Football


Figure Captions

Figure 1: Within- and Between-stride CMD over the season for individuals, with group mean in bold

Figure 2: Standardized (z-scored) effects of wellness and load on CMD

Table Captions

Table 1: Descriptive statistics for the 63 American Football athletes
Table 2: Linear Mixed Model Outputs
Table 3: Results from a generalized linear mixed model of flagged events