Gait & Posture 40 (2014) 521–525

Contents lists available at ScienceDirect

Gait & Posture

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

The use of regression and normalisation for the comparison of spatio-temporal gait data in children

P.C. Dixon [a], M.V. Bowtell [b], J. Stebbins [b,*]

[a] Department of Engineering Science, University of Oxford, Oxford, UK
[b] Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust, Oxford, UK

ARTICLE INFO

Article history:
Received 18 July 2012
Received in revised form 31 March 2014
Accepted 18 June 2014

Keywords:
Spatio-temporal
Children
Gait
Normalisation
Speed

ABSTRACT

Spatio-temporal parameters (STPs) are fundamental gait measures often used to compare children of different ages or gait ability. In the first case, non-dimensional normalisation (ND) of STPs using either leg-length or height is frequently conducted even though the process may not remove known inter-subject variability. STPs of children with and without disability can be compared through matched databases or using regression driven prediction. Unfortunately, database assignment is largely arbitrary and previous regressions have employed too few parameters to be successful. Therefore, the aims of this study were to test how well actual and ND STPs could be predicted from anthropometrics and speed and to assess if self-selected speed could be predicted from anthropometrics using multivariate regression in a cohort of eighty-nine typically developing children. Equations were validated on an extraneous dataset. We found that equations for actual step length, stride length, and cadence explained more than 84% of the variance compared to their ND counterparts. Moreover, only leg-length ND versions of these parameters were linearly proportional to speed. Prediction of single and double limb support times was weaker ($R^2 = 0.69$ and 0.72, respectively) and we were unable to predict self-selected speed ($R^2 < 0.16$) suggesting the use of anthropometrics is inappropriate for this purpose. Validation was successful for most STPs except in children lying near or outside the normal ranges and for gait speed. Clinically, regression could be used to quantify the difference between a patient's actual and theoretical STPs, allowing for monitoring of progress pre- and post intervention.

1. Introduction

Spatial-temporal parameters (STPs), such as self-selected speed, stride-length and cadence are basic measures of gait relating to foot-strike and foot-off placement and timing. There are differences in STPs between children with gait pathologies and their typically developing peers [1,2], adults and children [3,4], and amongst children of different ages. Amongst typically developing children, increased stride-length and decreased cadence lead to higher walking speeds with increasing age [5–8]. These differences can be attributed not only to anthropometric variability, such as leg-length or mass, but also to the neuromaturation effects of age [5]. Maturation plays a larger role in the early years whilst growth dominates the later stages of childhood [9–11].

STPs are often used to compare gait characteristics between children with varied anthropometrics or of different gait abilities, often walking at different speeds. There is therefore a requirement that comparisons account for the known variability between subjects. In the former case, the nondimensional normalisation (ND) approach of Hof [12] is often used and has been shown to effectively reduce inter-subject variability [13,14] and is used to compare subjects of different sizes, walking at similar ND speeds [15]. Yet the ND approach assumes proportional scaling and might not remove all age-related variability; in particular, variability arising from developmental differences may persist [5]. For the latter situation, STPs of children with and without gait pathology can be compared using large datasets grouped by age [6,8] or gait speed [16]; however, it remains unclear if grouping accounts for all the predictable variability in the subject groups. Alternatively, regression analysis may be used to predict expected gait parameters given no gait pathology and then determine how those with limited gait ability compare to their expected gait parameters. In a study by Stansfield et al. [15], ND STPs were regressed against ND speed only as maturation effects were assumed to be minimal in the cohort aged 7–12 years being

* Corresponding author at: Nuffield Orthopaedic Centre, Oxford University Hospitals NHS Trust, Oxford Gait Laboratory, Windmill Road, Headington, Oxford OX3 7HE, UK. Tel.: +44 1865227609; fax: +44 1865227949.
E-mail address: julie.stebbins@ouh.nhs.uk (J. Stebbins).

http://dx.doi.org/10.1016/j.gaitpost.2014.06.009
0966-6362/

investigated; however, the authors discouraged using their equations for predictive purposes. Possibly other factors, not included in the analysis, may have helped explain additional variability. An earlier study by the same lead author, investigating a slightly younger cohort, suggested that age might also be an important parameter [14]. It is conceivable that STPs could be best predicted by simultaneously analysing multiple anthropometric quantities such as height, leg-length, mass, and age as well as gait speed. In the aforementioned studies, gait speed was found to be a strong predictor of other STPs; yet, it remains unclear if self-selected speed itself can be reliably predicted from subject characteristics. In pathological populations this appears to be the case [17,18]; however, predictions in typically developing children are more rare. Vaughan et al. [5] reported that ND speed followed an exponential relationship with age: increasing rapidly in infants and reaching more stable values by age four. Perhaps additional anthropometric terms might further improve this relationship.

Therefore, the aims of this study were to test via multiple regression analysis whether (1) leg-length ND effectively removes the dependent relationship of speed, stride-length and cadence on anthropometrics and if (2) additional anthropometric terms could be used to generate better predictive equations for actual STPs. The first aim will allow critical evaluation of the use of leg-length normalised STPs for comparison of gait measures across different populations. The development of regression equations with strong predictive ability may improve the accuracy of comparison of gait data between children of different sizes, ages, and gait abilities.

2. Methods

2.1. Subjects

Fifty girls and forty-four boys (3–16 years) performing barefoot walking trials whilst fitted with the plug-in gait marker set [19] were extracted from our laboratory database (Table 1). Criteria for inclusion were: no known motor system pathology, walking independently, and experiencing no pain whilst walking. Ethical approval was granted by the local healthcare research ethics committee.

2.2. Data collection and processing

Six walking trials at self-selected speed were collected of each subject using 12 MX cameras and Nexus Software (Vicon, Oxford Metrics, Oxford, UK). A single representative trial for each subject was selected based on visual inspection of lower-limb joint kinematic and kinetic traces. Gait events were determined from the force plate data using a 10 N threshold, and verified visually from toe, heel, and ankle marker trajectories. Stride-length (m) and cadence (steps/min) were extracted using a Nexus plug-in (Parameter Calculator, Vaquita, Zaragoza, Spain) and defined as the ankle marker displacement in the direction of travel between consecutive foot strikes of the same foot and the number of steps per minute, respectively. The average of both legs was taken for each parameter. Finally, the open-source Biomechanical Toolkit [20] was used to import c3d files into Matlab (v2012b, The Mathworks, Inc., Natick, USA) where gait speed was computed by taking the average of the derivative of the sacral (SACR) marker position over a number of consecutive steps.

2.3. Normalisation

The ND was applied to the STPs according to Hof [12] using leg-length (anterior superior iliac spine to medial malleolus, via the medial femoral condyle) for stride-length. Similarly, cadence and speed were normalised to gravity and leg-length:

$$\| \mathrm{SL} \| = \frac{\mathrm{SL}}{\mathrm{LL}} \quad (1)$$

$$\| \mathrm{c} \| \; \mathrm{c} \times \sqrt{\frac{\mathrm{LL}}{g}} \quad (2)$$

$$\| \mathrm{v} \| \; \frac{\mathrm{v}}{\sqrt{g \times \mathrm{LL}}} \quad (3)$$

where SL: stride-length; c: cadence; v: speed; LL: leg-length; g: gravity (9.81 m/s^2).

2.4. Regression analysis

Multicollinearity between predictors (height, leg-length, body mass, age, and self-selected speed) was tested using the variance inflation factor (VIF) before performing stepwise multiple regression analysis [21]. Regression equations for stride-length and cadence were derived using leg-length, body mass, age and self-selected speed as predictor variables, whilst self-selected speed was regressed against these anthropometric quantities only. Both actual (raw) and ND forms for each STP were considered for the regression analysis. All models were executed via the LinearModel function running the stepwise option in the Matlab statistical toolbox (v2012b, The Mathworks, Inc., Natick, USA). The model only considered linear terms without interactions between variables. R^2 (for a single predictor) or adjusted R^2 (otherwise)

Table 1
Subject anthropometrics, self-selected speed, and STPs by age group.

Age (yrs)	Mass (kg)	Leg length (m)	Speed (m/s)	Stride length (m)	Cadence (steps/min)	ND speed	ND stride length	ND cadence
3–4	16.81 (15.17, 18.45)	0.49 (0.46, 0.52)	1.28 (1.13, 1.44)	0.86 (0.77, 0.94)	176.34 (16.621, 18.648)	0.59 (0.52, 0.65)	1.76 (1.57, 1.95)	0.66 (0.61, 0.70)
5–6	19.62 (18.14, 21.09)	0.57 (0.55, 0.60)	1.31 (1.22, 1.40)	0.99 (0.94, 1.04)	157.27 (148.52, 166.02)	0.55 (0.51, 0.60)	1.73 (1.65, 1.81)	0.63 (0.60, 0.66)
7–8	24.87 (23.58, 26.16)	0.66 (0.65, 0.68)	1.36 (1.32, 1.40)	1.10 (1.07, 1.12)	147.54 (144.80, 150.27)	0.53 (0.52, 0.55)	1.67 (1.62, 1.71)	0.64 (0.63, 0.65)
9–10	32.81 (31.48, 34.14)	0.73 (0.71, 0.75)	1.37 (1.31, 1.42)	1.20 (1.17, 1.24)	135.03 (133.01, 137.05)	0.51 (0.49, 0.52)	1.65 (1.62, 1.68)	0.61 (0.60, 0.63)
11–12	39.23 (36.96, 41.44)	0.81 (0.80, 0.84)	1.35 (1.31, 1.40)	1.28 (1.24, 1.33)	126.42 (123.41, 129.28)	0.48 (0.46, 0.50)	1.57 (1.53, 1.63)	0.60 (0.59, 0.62)
13–14	54.41 (50.34, 58.48)	0.87 (0.86, 0.89)	1.49 (1.43, 1.54)	1.43 (1.37, 1.48)	123.90 (121.05, 126.74)	0.51 (0.49, 0.52)	1.63 (1.59, 1.67)	0.62 (0.60, 0.63)
15–16	62.68 (60.34, 65.03)	0.90 (0.88, 0.92)	1.40 (1.36, 1.44)	1.43 (1.39, 1.47)	117.12 (114.40, 119.84)	0.47 (046, 049)	1.59 (1.56, 1.62)	0.59 (0.58, 0.60)

Mean (confidence interval).