Available online at www.sciencedirect.com
ScienceDirect
Gait & Posture 28 (2008) 323–326
GAIT & POSTURE
www.elsevier.com/locate/gaitpost

Reliability and validity of the Edinburgh Visual Gait Score for cerebral palsy when used by inexperienced observers

A.M.L. Ong [a], S.J. Hillman [b], J.E. Robb [c,*]

[a] College of Medicine and Veterinary Medicine, University of Edinburgh, Edinburgh, Scotland, United Kingdom
[b] National Centre for Training and Education in Prosthetics and Orthotics, University of Strathclyde, Glasgow, Scotland, United Kingdom
[c] Anderson Gait Analysis Laboratory, Royal Hospital for Sick Children, Department of Orthopaedic Surgery, Sciennes Road, EH9 1LF Edinburgh, Scotland, United Kingdom

Received 4 December 2007; received in revised form 19 January 2008; accepted 19 January 2008

Abstract

The Edinburgh Visual Gait Score (EVGS) for cerebral palsy has been validated for observer reliability and validity for observers experienced in gait analysis. This study investigated the reliability and validity of the EVGS for observers inexperienced in gait analysis. Six medical students used the score to analyse videotapes from the original study by Read et al. [Read HS, Hazlewood ME, Hillman SJ, Prescott RJ, Robb JE. Edinburgh visual gait score for use in cerebral palsy. J Pediatr Orthop 2003;23:296–301]. These were viewed on two separate occasions to provide inter- and intra-observer reliability, and the results of the numerical items were compared to those from three-dimensional (3D) gait analyses for validity. Observer agreement was tested using Coefficient of Repeatability (CoR), percentage of complete agreement and the kappa statistic. The CoR for inter-observer agreement for inexperienced observers was 5.99/5.07 (Session 1/Session 2) compared to 4.60/3.95 (Session 1/Session 2) for experienced observers. The CoR for intra-observer agreement for inexperienced observers was 5.15 compared to 4.21 for experienced observers. There was complete agreement for 52% of the 10 numerical items with 3D-gait analysis data for inexperienced observers compared to 64% for experienced observers. Ranking of reliability of individual items was similar between the two groups and was generally best for events occurring at the foot and ankle. Observations of gait events by the inexperienced observers using the EVGS were reasonably reliable but not very accurate when compared to experienced observers and 3D-gait analysis.

Keywords: Gait; Cerebral palsy; Visual assessment

1. Introduction

Although three-dimensional gait analysis (3DGA) has become widely accepted as the gold standard for evaluating complex locomotor disorders it is not widely available worldwide. In 2003 Read et al. [1] produced the Edinburgh Visual Gait Score (EVGS) for cerebral palsy using experienced observers. The Score showed good intra-observer and inter-observer reliability and almost two-thirds of the observations correlated with instrumented 3DGA data, signifying good validity. Subsequently the Score was compared to other measures (Gillette Gait Index (GGI), Gillette Functional Assessment Questionnaire and dimensionless speed) and showed significant correlations, especially with the GGI [2].

One problem in studies of gait is defining observer experience and how to quantify this. The purpose of this study was to investigate intra- and inter-observer reliability and validity of the EVGS when used by observers who were not specifically experienced in clinical gait analysis. Our hypotheses were firstly that the inexperienced observers would be less reliable than experienced observers and secondly that there would be a discernible learning effect and improvement of scores between the first and second viewings.

* Corresponding author. Tel.: +44 131 536 0834; fax: +44 131 536 0852.
E-mail address: j.e.robb@btinternet.com (J.E. Robb).

0966-6362/$ – see front matter
doi:10.1016/j.gaitpost.2008.01.008

2. Materials and methods

2.1. Subjects

The same video recordings of the five subjects used in Read et al's. [1] study were viewed by six 4th year undergraduate medical students. These observers had no expertise in the visual assessment of gait or cerebral palsy management but all had the same exposure to undergraduate locomotor system and neurology teaching. To ensure consistency all six received a standardised briefing about normal gait kinematics and an explanation of how to use the EVGS. The observers then used the EVGS to assess each limb of the subjects on the video recordings on two occasions, which were separated by an interval of 2 weeks. The subjects on the videotapes were assessed in a different order at the second viewing. The observers were allowed to view the videotape in slow motion or freeze-frame and no limit was placed on the time taken to view the subjects' gait.

2.2. Statistical analyses

The results from the two sessions were compared to calculate intra-observer reliability, using the Bland and Altman's Coefficient of Repeatability (CoR) [3]. Inter-observer reliability for the 17 items of the EVGS was examined in two ways: the percentage of complete agreement between observers and the kappa statistic. Grading of the kappa scores were carried out according to Landis and Koch [4].

The results of the 10 numerical items from the EVGS were compared to 3DGA data, considered to be the gold standard, to validate the observers' scores. Complete agreement was noted if the joint angle ranges scored by the observers were within the defined limits of the 3DGA values. The percentage of complete agreement was then calculated for each of the ten numerical items. These results from inexperienced observers were then compared with the results of the experienced observers from Read et al. [1].

3. Results

3.1. Intra-observer reliability

The mean CoR for intra-observer agreement for the inexperienced observers (1–6) was 5.15 (range 3.79–7.27) compared to a mean CoR of 4.21 (range 3.29–5.93) for experienced observers (A–E) from the study of Read et al. [1] (Table 1).

Table 1
Intra-observer reliability

Inexperienced observers	CoR	Experienced observers	CoR
Observer 1	4.90	Observer A	3.63
Observer 2	7.27	Observer B	4.20
Observer 3	5.93	Observer C	5.93
Observer 4	4.77	Observer D	3.29
Observer 5	3.79	Observer E	4.00
Observer 6	4.24		
Mean	5.15	Mean	4.21

3.2. Inter-observer reliability

The CoR for inter-observer agreement for inexperienced observers was 5.99/5.07 (Session 1/Session 2) compared to 4.60/3.95 (Session 1/Session 2) for experienced observers.

3.3. Inter-observer reliability for individual items of the EVGS

Inexperienced observers were generally less reliable than experienced observers for the 17 items in the EVGS. However, the ranking of reliability of individual items was similar between the two groups and was generally best for events occurring at the foot and ankle (Table 2). The most reliable item was initial contact which, for inexperienced observers, showed a kappa of 0.69 and percentage agreement of 80%, while experienced observers showed a kappa of 0.94, and percentage of complete agreement of 96% (Table 3). A comparison of inter-observer reliability for individual items between inexperienced and experienced observers using kappa scores is shown in Table 2. The comparison of inter-observer reliability for individual items between inexperienced and experienced observers using percentage of complete agreement is shown in Table 3.

3.4. Validity with 3DGA

There was complete agreement for 52% of the 10 numerical items with 3DGA data for inexperienced observers compared to 64% for experienced observers (Table 4). The inexperienced observers showed similar levels of accuracy, ranging from 46.5% complete agreement for the least accurate observer to 55.5% for the most accurate observer. There was also no notable difference in accuracy between the first and second sessions, with 50.5% mean accuracy for the first session and 53.3% mean accuracy for the second session.

4. Discussion

Inexperienced observers achieved moderate intra-observer reliability (mean CoR 5.15) compared with a mean of CoR 4.21 for experienced observers. The CoR represents the range within which 95% of the differences between two measures of the same quantity will be expected to lie. Thus a difference of less than or equal to approximately five would be expected when the same inexperienced observer, using the EVGS, repeats an observation of a patient on two separate occasions. The CoR of inexperienced Observer 2 was much higher than those for the other observers. If Observer 2's results were considered as an outlier and excluded, the CoR of the inexperienced observers would be similar to those of the experienced observers, signifying comparable reliability. The reason why this observer was less reliable is not clear but the CoR for the remaining