

Concurrent Validity Between Live and Home Video Observations Using the Alberta Infant Motor Scale

Marike Boonzaaijer, PT, MSc; Ellen van Dam, PT, MSc; Ingrid C. van Haastert, PT, PhD; Jacqueline Nuysink, PT, PhD

Department of Physiotherapy (Ms Boonzaaijer and Dr Nuysink), Institute for Human Movement Studies, Research Center for Innovation in Healthcare, University of Applied Sciences Utrecht, the Netherlands; Department of Neonatology (Dr van Haastert), Wilhelmina Children's Hospital, University Medical Center Utrecht, the Netherlands; and Department of Rehabilitation (Ms van Dam), Academic Medical Centre, University of Amsterdam, the Netherlands.

Purpose: Serial assessment of gross motor development of infants at risk is an established procedure in neonatal follow-up clinics. Assessments based on home video recordings could be a relevant addition.

Methods: In 48 infants (1.5-19 months), the concurrent validity of 2 applications was examined using the Alberta Infant Motor Scale: (1) a home video made by parents and (2) simultaneous observation on-site by a pediatric physical therapist. Parents' experiences were explored using a questionnaire.

Results: The intraclass correlation coefficient agreement between live and home video assessment was 0.99, with a standard error of measurement of 1.41 items. Intra- and interrater reliability: intraclass correlation coefficients were more than 0.99. According to 94% of the parents, recording their infant's movement repertoire was easy to perform.

Conclusion: Assessing the Alberta Infant Motor Scale based on home video recordings is comparable to assessment by live observation. The video method is a promising application that can be used with low burden for parents and infants. (*Pediatr Phys Ther* 2017;29:146–151)

Key words: gross motor development, home video recording, infant, motor assessment, validation study

INTRODUCTION

Screening gross motor development of infants to detect delays is a common practice for physical therapists who

are developmental specialists. The Alberta Infant Motor Scale (AIMS) is a well-known tool to assess gross motor performance in early infancy.¹ However, the accuracy is reduced if there is only 1 assessment.²⁻⁴ The assumption that the sequence and rate of gross motor development is stable within a child has been questioned.^{2,5-7} More evidence on inter- and intraindividual variability of gross motor development in infants is needed.^{2,3,8} However, longitudinal research is time consuming and testing in an outpatient setting can be burdensome for parents and infants. If the test is preplanned, there is no guarantee that the state of the infant at that testing will result in a valid assessment. For these reasons, a research project was designed: the Gross mOTor Development of Infants using home Video registration with the Alberta Infant Motor Scale (GODIVA). Parents were invited to make a structured video of their infants' gross motor repertoire in their home environment.

The reliability and validity of the AIMS are good to excellent.⁹ However, applying the AIMS in a home video setting requires validation.¹⁰ The main purpose of this study was to assess the concurrent validity between the AIMS score based on live observation (established procedure) and the AIMS score based on home video observation (new procedure). We hypothesized that the AIMS score obtained via home video registration is comparable to the score obtained by live observation. We examined the intra- and interrater reliability of the video method in the study. Feasibility of the video method for parents was explored.

0898-5669/292-0146

Pediatric Physical Therapy

Copyright © 2017 The Authors. Published by Wolters Kluwer Health, Inc. on behalf of Wolters Kluwer Health, Inc. and the Academy of Pediatric Physical Therapy of APTA

Correspondence: Marike Boonzaaijer, PT, MSc, Research Centre for Innovation in Healthcare, University of Applied Sciences Utrecht, PO Box 12011, 3501 AA, Utrecht, the Netherlands (marike.boonzaaijer@hu.nl).

Grant Support: This study was funded by Foundation Innovation Alliance—Regional Attention and Action for Knowledge circulation (SIA RAAK), project number 2013-53P, and the Netherlands Organization for Scientific Research, project number 023.006.070, the Netherlands.

This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

Supplemental digital content is available for this article. Direct URL citation appears in the printed text and is provided in the HTML and PDF versions of this article on the journal's Web site (www.pedpt.com).

The authors declare no conflicts of interest.

DOI: 10.1097/PEP.0000000000000363

METHODS

Design

A validation study design was used to determine the concurrent validity of the new and the original method. Parents were asked to complete a digital questionnaire that included questions on the feasibility of the video method. The study was approved by the Medical Ethical Committee of the University Medical Centre Utrecht, the Netherlands.

Participants

Infants (<19 months old) and parents were recruited from April through October 2014 using convenience sampling. Participation was open for parents who were interested in the study (eg, recruited at birth centers and well-baby clinics) or with a question on the motor development of their infant (recruited at physical therapy practices and included before intervention). Parents had to understand Dutch. Infants with known abnormal movement patterns were excluded. If abnormal movement patterns were observed at the video registration, parents and the family doctor would be informed and the infant was excluded from the study. Parents with a professional background being a physical therapist were excluded because of their knowledge on motor development. Both parents provided written informed consent.

Assessment Tool

Gross motor development was assessed using the AIMS, which was designed to evaluate the gross motor maturity of infants from birth to independent walking.¹ The original normative values were based on data from 2202 infants born in Alberta, Canada, and recently reevaluated.⁹ The scale contains 58 motor items divided into 4 subscales: prone (21 items), supine (9 items), sitting (12 items), and standing (16 items). Each item is described in detail considering the weight-bearing surface of the body, the posture necessary to achieve the gross motor skill, and the antigravity or voluntary movement of the infant. The total raw score can be converted into a percentile rank and/or z score. The reliability and content validity of the test are good.^{11,12}

Questionnaire

A digital questionnaire (Supplemental Digital Content 1, available at <http://links.lww.com/PPT/A137>) was developed by the researchers and consisted of 25 questions on a 5-point Likert scale. To characterize the study sample, questions were included about birth weight and gestational age of the infant. Parents were asked about their age, educational level, and knowledge of motor development. Questions on feasibility included technical and operational aspects of the recording.

PROCEDURES

The Video Method

Experienced pediatric physical therapists and researchers developed the method. To obtain videos suitable for rating the

gross motor performance, tutorial materials supported filming. The method consists of an instructional video and checklist (see Supplemental Digital Content 2 [<http://links.lww.com/PPT/A138>], 3 [<http://links.lww.com/PPT/A139>], and 4 [<http://links.lww.com/PPT/A140>]) for 3 age groups, which are adjusted to the motor abilities of the infant: group 1: 0 to 5.5 months, group 2: 5.5 to 8.5 months, and group 3: 8.5 to 19 months. Parents received the instructions that fit the motor abilities of their infants. Parents were allowed to use their mobile phone, tablet, or video camera. One parent recorded and 1 interacted with their infant. When only 1 parent was present, someone familiar was asked to film. The infant was undressed with the exception of a diaper and onesie. Filming was completed when parents captured the 4 different postures and movements. The recording was secured and stored at our research center. Parents received feedback on the motor performance of their child.

The Testers

Twelve pediatric physical therapists, familiar with the AIMS, attended 2 training sessions of 3 hours led by experts (ICvH, JN). Scoring gross motor performance of infants from videos was practiced and results were discussed using the AIMS administration guidelines. At the completion of training, each tester scored 2 video-recorded AIMS assessments. In order to be considered a tester, the therapist had to obtain a total raw score for each video within a range of ± 2 items compared with the consensus score set in a consensus meeting with 4 experts (ICvH, JN, EvD, MB) before the training. The 2-item range was derived from the acceptable range of the standard error of measurement (SEM) (1-2 items).¹

Home Video Recording and Assessment

The testers scheduled appointments with the parents at home. The parents made a home video recording while the tester observed the gross motor behavior of the infant. The parents were asked to use minimal infant handling. Motor behavior had to be spontaneous or elicited by presenting toys to the infant. Testers were told not to help parents make the video or handle the child. The “gold standard” in this study consisted of a live observation where the handling and prompting of the infant was completed by a pediatric physical therapist. Testers were allowed to complete extra observations or handling if necessary after parents completed the video recording.

A second tester rated the videos of motor behavior at the research center. Testers were masked to the AIMS scores of the other tester to ensure their ratings were independent and free from bias. The testers exchanged roles at random during the study.

Inter- and Intrater Reliability

Videos were used to evaluate the interrater and intrater reliability among 3 trained testers. They were masked with regard to the original scores. Each tester rescored 15 videos at random for a second time after at least 5 weeks from last scoring.

TABLE 1

Range of Age and Raw AIMS Scores in 3 Age Groups

Group	Sample Size	Male/Female	Age, wk Mean (SD); Range	Range AIMS Raw Scores (Live and Video)
1	16	6/10	16 (5.8); 4.9-25.6	3-6
2	12	6/6	30.3 (6.0); 22.7-42.6	17-31
3	20	12/8	54.2 (10.8); 31.7-78	32-58
Total	48	24/24	35.5 (18.7); 4.9-78	3-58

Abbreviations: AIMS, Alberta Infant Motor Scale; SD, standard deviation.

STATISTICAL METHODS

Concurrent Validity AIMS Video- and Live Observations and Reliability

The raw scores were used to determine the degree of agreement between the AIMS scores based on live observations and the AIMS scores based on the home video observations. High within-observer agreement is a prerequisite for obtaining valid scores. To analyze concurrent validity, intraclass correlation coefficients (ICCs) for a 3-way mixed effects model were used.¹³ The accepted level of agreement was 0.90.¹⁴ A Bland-Altman plot¹⁵ with Limits of Agreement was used to visualize the differences between the 2 measurements. To examine the measurement error in the 2 scores, the SEM was used and determined to be a maximum of 2 items.¹ The smallest detectable change (SDC) was calculated from the SEM. A 1-tailed *t* test was used to analyze the mean difference. The subscales of the AIMS were analyzed on the 3 age groups. The AIMS percentile scores¹ were used to investigate the number of cases for which the outcomes would be inconclusive.

Inter- and Intrarater Reliability AIMS Video Method

Because of a heterogeneous sample and expecting benefits of rescoring video material, we hypothesized that the ICC for

both the interrater and intrarater reliability would be at least as good as the reliability between the live and video assessments (ICC >0.90) and that the SEM, less than 2 items on the total raw score would be acceptable.¹ Statistical Package for the Social Sciences 21.0 (IBM SPSS Statistics for Windows, Version 21.0, Armonk, New York) was used for analyses.

RESULTS

Twelve testers completed the assessments. Videos of 52 infants, all of good technical quality, were obtained. Four videos (6%) were excluded because of failure to follow procedures. One infant was fully clothed during filming, 1 set of parents did not use the appropriate instructions during filming, and 2 videos were not made on 1 day.

The scores of 48 infants (24 males) were compared. The mean birthweight was 3432 g (range: 2500-4365 g). Infants were at least 37 weeks of gestation at birth and 1.5 to 18.5 months. The minimal total raw AIMS was 3 and the maximum 58 (Table 1).

Concurrent Validity AIMS Video and Live Observation

Figure 1 graphs the differences in AIMS scores between the live observation and the video method. The mean difference was

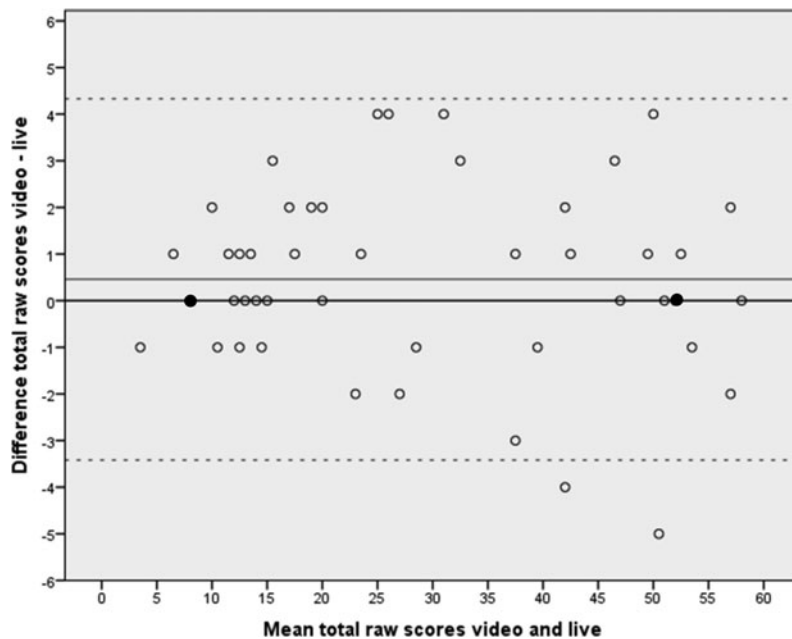


Fig 1. Bland-Altman plot: video-live (n = 48). Thick line (—) indicates mean difference total raw scores (0.46, standard deviation = 1.98); — shows absolute agreement (video score = live score); - - - shows limits of agreement (-3.42 vs +4.33; 95% of scores); and • represents 2 cases.

TABLE 2

Mean Differences Raw Scores in Subscales AIMS

Subscale Item AIMS	Sample Size	Male/Female	Mean Difference Subscale Video—Live (SD)	Range Total Raw Scores (Live and Video)	Difference In Raw Score Video—Live (SD)
Prone (21 items)	48	24/24	0.13 (0.56)	1-21	0.25 (1.1)
Supine (9 items)	48	24/24	0.10 (0.42)	1-9	0.21 (0.85)
Sitting (12 items)	48	24/24	-0.02 (0.33)	0-12	-0.04 (0.65)
Standing (16 items)	48	24/24	0.02 (0.42)	1-16	0.04 (0.85)
Total (58 items)	48	24/24	0.46 (1.98)	3-58	0.46 (1.98)

Abbreviations: AIMS, Alberta Infant Motor Scale; SD, standard deviation.

0.46 (standard deviation = 1.98), but not significant ($P = .115$; 95% confidence interval [CI] = -0.116 to $+1.033$; Table 2, Table 3). In 12 cases there was absolute agreement, in 23 cases the video observation was rated higher (score difference > 0 , mean difference [MD] 2.04 items, minimum 1 to maximum 4 items), and in 13 cases the live observation was rated higher (score difference < 0 , MD 1.92 items, minimum 1 to maximum 5 items). In 5 cases, there were considerable score differences: in 4 cases 4 items and in 1 case 5 items score differences between the 2 observations. In a few cases the clinical outcomes on the 2 assessments would lead to different advice to parents. Using the $p5$ as cutoff point,¹ the percentile ranks for 3 cases the infant (1, 5, and 8 months old) scored below the $p5$ in 1 assessment and above the $p5$ in the other assessment.

The ICC between the scores obtained by live and video observation was 0.99. The lowest ICC was in age group 2 (0.89) (Table 3). To determine absolute agreement given in items of the test, the SEM was calculated to be 1.41. In age group 3, the highest SEM was 1.63 and the smallest value of 0.80 was in age group 1. The SDC was calculated from the SEM¹⁰ and was 3.88 items. This is the minimal amount of change that must be observed before the change can be considered to exceed the variation and measurement error at the 95% CI.

Table 4 has the ICC, SEM, and SDC of the 4 subscales of the AIMS. The ICC in supine position is lowest (0.94 item). The SEM is highest in prone position (0.79 item) as is the SDC (2.19 items). This subscale consists of 21 items, the largest amount of the 4 subscales.

TABLE 3

Validity Results in 3 Age Groups

Group	ICC (3-Way Mixed)	SEM	SDC	MD
1	0.94	0.80	2.20	0.31
2	0.89	1.54	4.27	1.25
3	0.95	1.63	4.50	0.10
Total	0.99	1.41	3.88	0.46 (standard deviation = 1.98)

Abbreviations: ICC, intraclass correlation coefficient; MD, mean difference; SDC, smallest detectable change; SEM, standard error of the measurement.

TABLE 4

Validity Results in Subscales AIMS

Subscales AIMS	ICC (3-Way Mixed)	SEM	SDC
Prone	0.99	0.79	2.19
Supine	0.94	0.59	1.64
Sitting	0.99	0.46	1.28
Standing	0.98	0.60	1.66
Total	0.99	1.41	3.88

Abbreviations: AIMS, Alberta Infant Motor Scale; ICC, intraclass correlation coefficient; SDC, smallest detectable change; SEM, standard error of the measurement.

Inter- and Intrater Reliability AIMS Video Method

An ICC of 0.99 on the total raw scores between 3 testers indicates a high interrater reliability of the video method, the average SEM of 0.92 item on the total raw score of the AIMS. The SDC was calculated to be 2.55 items. The intrater reliability of the video method had an ICC on the total raw score of 0.997 (range: 0.995-0.998). The SEM was 0.96 item and the SDC 2.66 items.

Feasibility

Fifty-one questionnaires were completed by the parents (86% by mothers). Almost 75% of the study sample had advanced education. Mean total time for completing instructions and filming was 36.4 minutes (standard deviation = 21.33; range: 5-90). Seventy-eight percent of the parents reported that their child demonstrated optimal motor performance or had new motor behavior. According to 94% of the parents, recording their infant's movement repertoire was easy to perform. Ten percent of the parents had doubts about sending a video of their child to professionals. In 96% of the cases, parents reported that making a home video was easy.

DISCUSSION

The results of this study support high degrees of agreement between an assessment based on a video recording made by parents and a simultaneous live assessment of the gross motor repertoire of an infant. The reliability of the video-recording

method itself was evaluated as good; both inter- and intrarater reliability had large correlations. The conclusions on the feasibility of the video method for parents are positive.

One of the most important findings in this study is the lack of a systematic difference in the total raw score between the video and the live observation scores, and no difference in the 4 subscales or in the 3 age groups. Scores obtained through video assessments were in general slightly higher than the live assessments (+0.46 item). In age group 2, ICC is the lowest (0.89 item) whereas the MD scores are the highest (1.25 item). This finding does not correspond with the ICCs that were found in the reliability study of the original AIMS,¹ where correlations were lowest in the youngest and oldest group of infants who performed fewer items. The lower correlation in age group 2, in the present study, is likely the result of a small sample ($n = 12$). The ICC of the subscale supine is slightly lower (0.94 item) than the other subscales. This might be due to the fact that this subscale consists of only 9 items.

Because there are no guidelines for an acceptable SEM, it has to be defined a priori according to the unit and purpose of the measurement. Before the study, a clinically acceptable SEM for the AIMS was set at 1 to 2 items. A SEM of 1.41 items meets this criterion. In the reliability study of the original AIMS,¹ the SEM was found to be 1.01 on the interrater reliability with 2 trained testers being present at 1 occasion, where the primary assessor was administering the test and the other was observing. The interrater reliability, in the present study, compared the live and video observation made by parents and rated by different testers resulted in an SEM of 1.41 items.

Because the SEM includes both method variation and between-rater variation, 1 of the main issues in this study was to establish the source of the error variance when there were differences found between the 2 scores. Were they due to the between-rater variation or to limitations of the video recording method? In 2 of the 6 cases when differences are 4 items or more, the live observer rated the infant respectively 4 and 5 items higher than the video observer did. In these 2 cases this was the result of more handling done by the live tester after the parents completed filming. However, because the video assessment scores were in general *higher* than the live assessment scores, we concluded that differences between the live and video scores in most cases have to be allocated to moderate reliability caused by the involvement of a large number of testers.

With an SDC of 3.88 items, an infant must show a progress of 4 items or more on the AIMS on the following assessment before it can be seen as a real change (95% CI), not the result of measurement errors. In clinical use of the AIMS, we expect this SDC not to be a limitation. It means a progress of, for instance, one item in each of the 4 subscales. The AIMS has been described to be sensitive to small increments of change over brief periods, even as short as a week.¹ Given the frequency of assessing gross motor development in a clinical setting, it can be expected that the detectable change in a next assessment will be greater than the measurement error.

In the design of the present study, the method of live observation and scoring the AIMS was considered to be the “gold standard” because it is an established procedure. By analyzing the data, it was not always possible to establish which score

(live or video) was the best representation of the actual gross motor performance of the infant. In some cases the live observer observed more items, but in other cases the live observer failed to observe items, which were present at the home video. Therefore, the “gold standard” assumption must be questioned, which means that the outcomes on validity should be interpreted with some caution.

The 12 pediatric physical therapists from the field, who obtained the data, were very diverse in age and years of experience. Making use of this large and heterogeneous group of testers added to the error variance but gave more insight into the potential use of the video method in clinical practice.

The high levels of reliability between and within testers indicate that the 3 trained testers can replicate their scores on the AIMS video method with accuracy. The SEM and SDC of the video method are lower (0.92 and 2.66 items, respectively) than those of the live and video method combined (1.41 and 3.88 items, respectively). This is an expected consequence of the involvement of fewer testers (12 vs 3 testers) and assessing only the video material. The findings on the home video method correspond to other reliability studies of the AIMS using video materials.¹⁶⁻¹⁸

Our study supports that in most cases parents are capable of making suitable videos that can be used to record a valid assessment of the gross motor behavior of their child. Asking parents to make a video that is used for assessment is relatively new. The video recording method depends partially on an adequate understanding of what and how to film. In recent research articles, there is evidence of parents being able to provide valid reports on early motor development of their children.^{12,19} This supports that parents have valid ideas about the gross motor development of their children. However, the educational levels of the parents could have positively influenced the quality of the video recordings. Further research is needed to make clear whether the video-recording method is feasible for parents of different social, ethnic, educational, and economic backgrounds. The feasibility of the video-recording method for parents who have an infant at risk (eg, prematurity) should be explored in future research.

This study also raises another important question: What is the best way to observe early gross motor performance? A live observation is not lasting. Retrospective scoring on the recollection of the observation can be liable to errors. More and more assessors who observe gross motor performance are using video recordings to improve the objectivity of the observation or test.²⁰⁻²² For instance, the agreement between video recordings and live assessments of the gross motor function measurement in children with cerebral palsy can be reliably scored using video recordings.²² A possible disadvantage, however, might be that professionals can only explore the motor performance shown in the video, which may provide incomplete information or interpretation.²³

Clinically, the video-recording method might become a promising addition to the established procedures of monitoring and assessing infants at risk. A key future application of the video-recording method could be in longitudinal research projects to develop infant gross motor trajectories. Repeated examiner-administered assessments in longitudinal studies are

expensive³ and can be burdensome for infants and parents. To make this home video method available for professionals, work must be done to realize a secured web-based design, which enables parents and professionals to interchange videos and feedback.

Another opportunity to use this method is teleconsultancy. Parents who live in rural areas and have concerns about the gross motor development of their infants but are not able to visit a hospital or physical therapy practice can use this home video method. After uploading their video recording onto a safe server, a trained pediatric physical therapist can assess the movement repertoire of the infant and, if needed, give practical advice or refer to a specialist.

The results of this study indicate that the AIMS home video method provides reliable and valid measurements that are interchangeable with the live assessments of the AIMS. However, parents must follow video procedures to obtain a valid measurement of the gross motor maturity of their child and therapists must use the precise descriptions for scoring the AIMS. This method allows parents to choose a suitable time for filming, so the infant can show the best motor performance in a home environment. Time and distance become less important barriers.

ACKNOWLEDGMENTS

We thank the testers for their live and video observations and the recruitment of children, Arine Sneep, Floortje Roest, and Nienke Reumer for their work on the inter- and intrarater reliability, and finally, all the infants and their parents for their enthusiastic participation.

REFERENCES

1. Piper MC, Darrah J, ed. *Motor Assessment of the Developing Infant*. 1st ed. Philadelphia, PA: Saunders; 1994.
2. Darrah J, Hodge M, Magill-Evans J, Kumbhani G. Stability of serial assessments of motor and communication abilities in typically developing infants—implications for screening. *Early Hum Dev*. 2003;72(2):97-110.
3. Adolph KE, Robinson SR, Young JW, Gill-Alvarez F. What is the shape of developmental change? *Psychol Rev*. 2008;115(3):527.
4. Roze E, Meijer L, Van Braeckel KN, Ruiters SA, Bruggink JL, Bos AF. Developmental trajectories from birth to school age in healthy term-born children. *Am Acad Pediatrics*. 2010;126(5):1134-1142.
5. Darrah J. Intra-individual stability of rate of gross motor development in full-term infants. *Early Hum Dev*. 1998;52(2):169-179.
6. Piek JP. The role of variability in early motor development. *Infant Behav Dev*. 2002;25(4):452-465.
7. Janssen AJ, Akkermans RP, Steiner K, et al. Unstable longitudinal motor performance in preterm infants from 6 to 24 months on the Bayley Scales of Infant Development—Second edition. *Res Dev Disabil*. 2011;32(5):1902-1909.
8. Piek JP, Dawson L, Smith LM, Gasson N. The role of early fine and gross motor development on later motor and cognitive ability. *Hum Mov Sci*. 2008;27(5):668-681.
9. Darrah J, Bartlett D, Maguire TO, Avison WR, Lacaze-Masmonteil T. Have infant gross motor abilities changed in 20 years? A re-evaluation of the Alberta Infant Motor Scale normative values. *Dev Med Child Neurol*. 2014;56(9):877-881.
10. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res*. 2010;19(4):539-549.
11. Piper MC, Pinnell LE, Darrah J, Maguire T, Byrne PJ. Construction and validation of the Alberta Infant Motor Scale (AIMS). *Can J Public Health*. 1992;83(suppl 2):46-50.
12. Libertus K, Landa RJ. The Early Motor Questionnaire (EMQ): a parental report measure of early motor development. *Infant Behav Dev*. 2013;36(4):833-842.
13. de Vet HC, Terwee CB, Knol DL, Bouter LM. When to use agreement versus reliability measures. *J Clin Epidemiol*. 2006;59(10):1033-1039.
14. Kottner J, Audigé L, Brorson S, et al. Guidelines for reporting reliability and agreement studies (GRRAS) were proposed. *J Clin Epidemiol*. 2011;64(1):96-106.
15. Bland JM, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986;327(8476):307-310.
16. Jeng S, Yau KT, Chen L, Hsiao S. Alberta Infant Motor Scale: reliability and validity when used on preterm infants in Taiwan. *Phys Ther*. 2000;80(2):168-178.
17. Valentini NC, Saccani R. Brazilian validation of the Alberta Infant Motor Scale. *Phys Ther*. 2012;92(3):440-447.
18. Syrengelas D, Siahianidou T, Kourlaba G, Kleisiouni P, Bakoula C, Chrousos GP. Standardization of the Alberta Infant Motor Scale in full-term Greek infants: preliminary results. *Early Hum Dev*. 2010;86(4):245-249.
19. Bodnarchuk JL, Eaton WO. Can parent reports be trusted? Validity of daily checklists of gross motor milestone attainment. *J Appl Dev Psychol*. 2004;25(4):481-490.
20. Wiles C, Newcombe R, Fuller K, Jones A, Price M. Use of videotape to assess mobility in a controlled randomized crossover trial of physiotherapy in chronic multiple sclerosis. *Clin Rehabil*. 2003;17(3):256-263.
21. Adde L, Helbostad J, Jensenius AR, Langaas M, Stoen R. Identification of fidgety movements and prediction of CP by the use of computer-based video analysis is more accurate when based on two video recordings. *Physiother Theory and Pract*. 2013;29(6):469-475.
22. Franki I, Van den Broeck C, De Cat J, Molenaers G, Vanderstraeten G, Desloovere K. A study of whether video scoring is a reliable option for blinded scoring of the Gross Motor Function Measure-88. *Clin Rehabil*. 2015; 29(8):809-815.
23. Fyfe S, Downs J, McIlroy O, et al. Development of a video-based evaluation tool in Rett syndrome. *J Autism Dev Disord*. 2007;37(9):1636-1646.